



**Raquel
da
Fontoura
Nicolette**

**Modelos Autorregressivos por Limiares em Séries
de Contagem**



**Raquel
da
Fontoura
Nicolette**

Modelos Autorregressivos por Limiares em Séries de Contagem

Dissertação apresentada às Universidades de Aveiro e Minho para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Matemática e Aplicações, realizada sob a orientação científica da Doutora Isabel Maria Simões Pereira, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro e sob coorientação científica do Doutor Manuel González Scotto, Professor Auxiliar com agregação do Departamento de Matemática da Universidade de Aveiro.

O júri

presidente

Doutor José Rodrigues Ferreira da Rocha

Professor Catedrático da Universidade de Aveiro (por delegação da Reitora da Universidade de Aveiro)

vogais

Doutora Maria de Nazaré Simões Quadros Mendes Lopes

Professora Catedrática, Faculdade de Ciências e Tecnologia, Universidade de Coimbra

Doutora Maria Eduarda da Rocha Pinto Augusto da Silva

Professora Associada, Faculdade de Economia, Universidade do Porto

Doutora Patrícia Cortés de Zea Bermudez

Professora Auxiliar, Faculdade de Ciências, Universidade de Lisboa

Doutora Isabel Maria Simões Pereira

Professora Auxiliar, Universidade de Aveiro (orientadora)

Doutora Maria da Conceição Soares Serra

Professora Auxiliar, Escola de Ciências, Universidade do Minho

Dedico este trabalho à Laís e ao Nei pelo apoio e por sempre ansiosos esperarem minha volta.

agradecimentos

Aos meus orientadores, Professora Isabel Pereira e Professor Manuel Scotto por sua gigantesca paciência e troca de experiências ao longo destes 5 anos.

Aos meus bons amigos feitos em Portugal, passageiros ou que permanecerão na minha vida. Agradeço por sua dedicação, carinho e boas risadas mesmo nas horas de desespero.

Agradeço aos meus eternos amigos no Brasil os quais me apoiaram todos os momentos para embarcar nesta aventura *d'além mar*.

Àqueles da minha família que me deram suporte emocional para ficar tanto tempo longe, mesmo privados da convivência física diária o mundo virtual nos possibilitou permanecer tão perto quanto nunca.

Agradeço a *CapesNei* pelo suporte financeiro adicional para poder concluir a tese. Sem tal financiamento nada disso hoje seria realidade.

E o agradecimento especial à Laís e ao Nei, meus maiores tesouros. Agradeço por não terem me esquecido nestes anos longe. Agradeço por me fazerem sorrir nos dias cinzentos de Aveiro. Agradeço por me sentirem presentes mesmo que distante. Agradeço por esperarem minha volta ansiosos.

Gostaria de expressar meu sincero agradecimento pelo suporte financeiro recebido pela Fundação Portuguesa para Ciência e Tecnologia (FCT) na forma da bolsa de doutoramento SFRH/BD/33385/2008.



palavras-chave

processos autorregressivos, valores inteiros, modelos definidos por limiares, modelos periódicos, reversible jump

resumo

A modelação e análise de séries temporais de valores inteiros tem sido alvo de grande investigação e desenvolvimento nos últimos anos, com aplicações várias em diversas áreas da ciência. Nesta tese a atenção centrar-se-á no estudo na classe de modelos baseados no operador *thinning* binomial.

Tendo como base o operador *thinning* binomial, esta tese focou-se na construção e estudo de modelos $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ e $\text{PSETINAR}(2; 1, 1)_T$, modelos autorregressivos de valores inteiros com limiares autoinduzidos e dois regimes, admitindo que as inovações formam uma sucessão de variáveis independentes com distribuição de Poisson.

Relativamente ao primeiro modelo analisado, o modelo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$, além do estudo das suas propriedades probabilísticas e de métodos, clássicos e bayesianos, para estimar os parâmetros, analisou-se a questão da seleção das ordens, no caso de elas serem desconhecidas. Com este objetivo consideraram-se algoritmos de Monte Carlo via cadeias de Markov, em particular o algoritmo *Reversible Jump*, abordando-se também o problema da seleção de modelos, usando metodologias clássica e bayesiana. Complementou-se a análise através de um estudo de simulação e uma aplicação a dois conjuntos de dados reais.

O modelo $\text{PSETINAR}(2; 1, 1)_T$ proposto, é também um modelo autorregressivo com limiares autoinduzidos e dois regimes, de ordem unitária em cada um deles, mas apresentando uma estrutura periódica. Estudaram-se as suas propriedades probabilísticas, analisaram-se os problemas de inferência e predição de futuras observações e realizaram-se estudos de simulação.

palavras-chave

autoregressive processes, integer values, threshold models, periodic models, reversible jump

abstract

The analysis of integer-valued time series has become an important area of research in the last two decades partially because its wide applicability in many areas of science. Arguably, models for integer-valued time series can be classified into two groups, namely integer-valued autoregressive models based on thinning operators and autoregressive conditional Poisson models. In this thesis, we restrict our attention to models belonging to the first class.

In particular, in this work self-exciting threshold integer-valued autoregressive models with two regimes and orders $p^{(1)}$ and $p^{(2)}$ (SETINAR(2; $p^{(1)}, p^{(2)}$), in short) and periodic SETINAR models of orders one and one (PSETINAR(2; 1, 1)_T, in short) are introduced and studied in detail.

Concerning the SETINAR(2; $p^{(1)}, p^{(2)}$) model, basic probabilistics and statistical properties of the models are analysed. Furthermore, the problem of model order selection is also addressed.

In what concerns the PSETINAR(2; 1, 1)_T model, issues related with the existence and uniqueness of a periodically stationary and causal PSETINAR(2; 1, 1)_T process are discussed. In addition to that parameter estimation and prediction are also covered.

Sumário

Lista de Abreviaturas	ix
Introdução	1
1 Fundamentos Teóricos e Estado da Arte	7
1.1 Operador <i>Thinning</i> Binomial	7
1.2 Modelos Autorregressivos de Valores Inteiros (INAR)	10
1.2.1 Modelo INAR(1)	10
1.2.2 Modelo INAR(p)	12
1.2.3 Estado da Arte	15
1.3 Elementos Fundamentais de Estatística Bayesiana	17
1.4 Métodos de Monte Carlo via Cadeias de Markov	19
1.5 Algoritmos de Ampliação de Dados	23
1.5.1 Algoritmo <i>Data Augmentation</i>	24
1.5.2 Algoritmo Expectation-Maximization (EM)	25
1.5.3 Algoritmo <i>Monte Carlo Expectation Maximization</i>	27
1.5.4 Aplicações dos Algoritmo RJMCMC e EM em Séries de Valores Inteiros	28
2 Processos Autorregressivos de Valores Inteiros Definidos por Limites Autoinduzidos	31
2.1 Definição do Modelo SETINAR	32
2.1.1 Representação Vetorial do Modelo SETINAR	33
2.2 Propriedades Gerais	35
2.2.1 Estacionaridade Estrita	35
2.2.2 Propriedades Probabilísticas	35
2.3 Métodos de Estimação	41

2.3.1	Estimadores de Mínimos Quadrados Condicionais	41
2.3.2	Verosimilhança Condicional	45
2.3.3	Inferência bayesiana	50
2.4	Estudo de Simulação	61
2.4.1	Análise dos Resultados de Estimação	63
2.4.2	Seleção de Modelos	71
2.5	Aplicação a Dados Reais	79
2.5.1	Aplicação 1	80
2.5.2	Aplicação 2	83
2.6	Conclusões	88
3	Processos Autorregressivos de Valores Inteiros Definidos por Limi- ares Periódicos Autoinduzidos	89
3.1	O Modelo PSETINAR(2; 1, 1) _T	90
3.2	Métodos de Estimação	98
3.2.1	Estimadores de Mínimos Quadrados Condicionais	99
3.2.2	Estimadores de Máxima Verosimilhança Condicional	103
3.3	Estudo de Simulação	111
3.3.1	Análise dos Resultados da Estimação	113
3.4	Predição no modelo PSETINAR(2; 1, 1) _T	120
3.4.1	Predição Pontual	120
3.4.2	Estudo de Simulação	122
3.5	Conclusões	128
4	Considerações Finais	130
	Apêndices	133
A	Cálculos Auxiliares	133
A.1	Algoritmo EM de Monte Carlo	133
A.1.1	Distribuições Condicionais Completas	138
A.2	Cálculos para o Modelo PSETINAR	141
A.2.1	Derivadas Parciais de $\frac{p_j(b-1 a)}{p_j(b a)}$	141
A.2.2	Derivadas Parciais de Segunda Ordem do Logaritmo da Função de Verosimilhança	143
B	Funções Implementadas no Programa R	147
B.1	Funções Utilizadas para o Modelo SETINAR	147

C	Dados de Manchas Solares	152
C.1	Aplicação 1	152
3.2	Aplicação 2	153

Lista de Figuras

2.1	Realizações de dois modelos $\text{SETINAR}(2; p^{(1)}, p^{(2)})$	34
2.2	Realizações do modelo SETINAR	62
2.3	Viés das estimativas de $\theta = (0.2; 0.1, 0.1; 3; 7)$ do modelo M_1	66
2.4	Viés das estimativas de $\theta = (0.3; 0.2, 0.1; 3; 3)$ do modelo M_2	67
2.5	Viés das estimativas de $\theta = (0.2, 0.1; 0.1; 3; 3)$ do modelo M_3	69
2.6	Viés das estimativas de $\theta = (0.6, 0.3; 0.6; 4; 1)$ do modelo M_4	70
2.7	Modelo M_1	75
2.8	Modelo M_2	76
2.9	Modelo M_3	77
2.10	Modelo M_4	78
2.11	Mancha solar - 15/7/2002 - Royal Swedish Academy of Sciences . . .	81
2.12	Série temporal do número anual de manchas solares de Wolfer ano período 1770-1869	81
2.13	FAC da série de Boulder	82
2.14	Série temporal do número total de grupos de manchas solares no período de 1 de Janeiro de 1990 a 31 de Dezembro de 1991.	84
2.15	FAC da Série de Palehua	86
2.16	FAC da série de Boulder	87
3.1	Gráfico de algumas realizações do modelo PSETINAR	112
3.2	Viés das estimativas de $\theta = (0.2, 0.65; 0.2, 0.1; 7; 3)$ do modelo M_1 . . .	115
3.3	Viés das estimativas de $\theta = (0.2, 0.1; 0.2, 0.1; 7; 3)$ do modelo M_2 . . .	116
3.4	Viés das estimativas de $\theta = (0.65, 0.3; 0.6, 0.35; 6; 6)$ do modelo M_3 . .	118
3.5	Viés das estimativas de $\theta = (0.8, 0.1; 0.65, 0.2; 7; 3)$ do modelo M_4 . .	119
3.6	Valores de x_{t+1} (preto), \hat{x}_{t+1} (vermelho), \hat{m}_{t+1} (verde) e $\hat{m}o_{t+1}$ (azul) de uma realização de M1 e M2	125

3.7	Valores de x_{t+1} (preto), \hat{x}_{t+1} (vermelho), \hat{m}_{t+1} (verde) e \hat{mo}_{t+1} (azul) de uma realização de M3 e M4	126
-----	---	-----

Lista de Tabelas

1.1	Propriedades do operador <i>thinning</i> binomial	8
1.2	Propriedades do operador <i>thinning</i> binomial	9
1.3	Propriedades do operador <i>thinning</i> binomial vetorial	10
1.4	Propriedades do modelo INAR(1)	11
2.1	Parâmetros dos modelos SETINAR(2; $p^{(1)}, p^{(2)}$)	61
2.2	Média amostral e EQM (em parêntesis) das estimativas dos parâmetros para os modelos M_1 e M_2	65
2.3	Média amostral e EQM (em parêntesis) das estimativas dos parâmetros para os modelos M_3 e M_4	68
2.4	Seleção de modelos pelo critério BIC	72
2.5	Seleção de modelos pelo critério DIC	73
2.6	Seleção de modelos usando o algoritmo RJMCMC	73
2.7	Seleção de modelos usando BIC, DIC e RJMCMC	74
2.8	Seleção do modelo, usando BIC, DIC e RJMCMC	83
2.9	Estimativas de θ para a série de Wölfer	83
2.10	Seleção do modelo, usando BIC, DIC e RJMCMC	85
2.11	Média amostral das estimativas de θ para a série Palehua	85
2.12	Média amostral das estimativas de θ para a série Boulder	85
3.1	Parâmetros dos modelos PSETINAR(2; $p^{(1)}, p^{(2)}$)	111
3.2	Média amostral das estimativas de θ para os modelos M_1 e M_2	114
3.3	Média amostral das estimativas de θ para os modelos M_3 e M_4	117
3.4	Parâmetros dos modelos PSETINAR(2; $p^{(1)}, p^{(2)}$)	122
3.5	EQM, EAM e FPTN de 20 predições um passo à frente para M1, M2 e M3	124
3.6	EAPMde 20 predições um passo à frente para M1, M2 e M3	124
C.1	Número da manchas solares de Wölfer - anual	152

3.2	Número total de grupos de manchas solares no período de 1 de janeiro de 1990 a 31 de dezembro de 1991 - Centro nacional de dados geofísicos de Boulder (Colorado, EUA)	153
3.3	Número total de grupos de manchas solares no período de 1 de janeiro de 1990 a 31 de dezembro de 1991 - Palehua Solar observatory (Hawai, EUA)	153

Lista de abreviaturas

AR	Autorregressivo com suporte em \mathbb{R} .
BIC	Critério de Informação de Bayes .
DA	Data augmentation - ampliação de dados.
EAM	Erro absoluto médio.
EAPM	Erro absoluto percentual médio
EM	Expectation-Maximization.
EQM	Erro quadrático médio.
FAC	Função de autocorrelação.
FPTN	Função perda tudo ou nada.
i.i.d	Independente(s) e identicamente distribuída(s).
INAR	AutoRegressivo com suporte em \mathbb{Z}_0^+ .
MCEM -	Monte Carlo Expectation-Maximization.
MCMC	Método de Monte Carlo via Cadeias de Markov.
MH	Metropolis-Hastings.
MQC	Mínimos quadrados condicionais.
MVC	Máxima verosimilhança condicional.
PSETINAR	Autorregressivo periódico com limiares autoinduzidos com suporte em \mathbb{Z}_0^+ .
RJMCMC	Método de Monte Carlo via Cadeias de Markov com saltos reversíveis.
SETINAR	Autorregressivo com limiares autoinduzidos com suporte em \mathbb{Z}_0^+ .
v.a.(’s)	variável(eis) aleatória(as).

Introdução

*“The most important questions of life are, for
the most part, really only problems of
probability.”*

Pierre-Simon Laplace

Nas últimas décadas o estudo de séries temporais de valores inteiros não negativos tornou-se mais presente em diversos ramos das ciências. O crescente interesse no estudo de séries de valores inteiros dá-se pelo fato da metodologia tradicional de estudo de séries temporais ficar limitada, na maioria das vezes, às séries de dados de valores reais.

Séries temporais de valores inteiros são obtidas através de processos de contagem definidos como o número de acontecimentos, objetos ou indivíduos em intervalos consecutivos de tempo. Como exemplos tem-se o número de nascimentos ou mortes num hospital a cada dia, o número de acidentes numa fábrica a cada mês, o número de viajantes de comboio a cada hora, o número de processos binários num período, entre inúmeras outras possibilidades.

No âmbito do estudo das séries temporais de valores inteiros, diversos trabalhos podem ser citados nas mais diferentes áreas das ciências tais como o turismo (Nordström, 1996; Brännäs *et al.*, 2002; Brännäs & Nordström, 2006), ciências do ambiente (Thyregod *et al.*, 1999; Scotto *et al.*, 2014), sistemas de fila de espera (Ahn

et al., 2000), biologia experimental (Zhou & Basawa, 2005), ciências sociais (McCabe & Martin, 2005), economia (Quoreshi, 2006; Brännäs & Quoreshi, 2010; Quoreshi, 2012), telecomunicações (Weiß, 2008b) e sistemas de alarme ótimos (Monteiro *et al.*, 2008).

A contagem discreta de eventos nos modelos usuais, devido à sua abordagem contínua, não é o procedimento mais adequado dado que o produto de uma constante real por uma variável aleatória de valor inteiro resulta numa variável aleatória real (Silva & Silva, 2003). Assim torna-se necessário o uso de uma metodologia adequada para o tratamento desse tipo de observações. Uma solução para este problema é substituir a multiplicação usual por uma operação aleatória discreta, chamada operação *thinning* binomial, definida por Steutel & van Harn (1979), a qual será detalhada posteriormente.

A partir da definição da operação *thinning* binomial, diversos modelos de contagem de dados utilizando esta operação têm sido propostos. Um desses modelos é o **INAR** (*INteger-valued AutoRegressive*) onde o processo imita a estrutura de correlação do conhecido processo linear Autorregressivo (AR).

Uma abordagem que é aliada aos modelos **INAR** no estudo de séries temporais é usar a Inferência bayesiana. Os métodos **MCMC** são uma ferramenta computacional importante para gerar amostras de distribuições de probabilidade não standardizadas e baseiam-se na construção de cadeias de Markov que apresentam como distribuição estacionária (de equilíbrio), a distribuição de probabilidade desejada. Para se implementarem os métodos **MCMC** há a necessidade de saber construir cadeias de Markov com distribuições de equilíbrio específicas. Metropolis *et al.* (1953) desenvolveram um algoritmo para esse fim, o qual foi posteriormente generalizado por Hastings (1970), sendo conhecido na literatura por *algoritmo de Metropolis-Hastings*. A associação destes algoritmos para a simulação de distribuições com o método de Monte Carlo para a aproximação de integrais conduz aos chamados métodos **MCMC** (ver, por exemplo, Gilks *et al.*, 1996). Um caso particular destes métodos é o *método de amostragem*

Gibbs, que foi introduzido por Geman & Geman (1984) para simular distribuições multivariadas, muito complexas e comuns em problemas de reconstrução de imagens. Gelfand & Smith (1990) mostraram como o algoritmo pode ser usado para simular distribuições *a posteriori*, e consequentemente a sua importância na resolução de problemas em Inferência bayesiana.

De uma forma geral não é muito difícil construir uma Cadeia de Markov com as propriedades desejadas. O maior problema consiste em determinar a partir de que iteração se pode considerar que a convergência para a distribuição de equilíbrio foi atingida. Assim o desafio consiste em estudar o problema teórico da convergência. Podem-se destacar, entre outros, os trabalhos de Tierney (1994), Roberts (1996), Besag & Green (1993) e Chan (1993). Os métodos de análise da convergência mais divulgados encontram-se na *package* CODA (Best *et al.*, 1995) e na *package* BOA (Smith, 2004), funcionando no ambiente R.

A velocidade de convergência está intimamente ligada com a escolha de distribuições proponentes que fazem parte integrante do algoritmo de Metropolis-Hastings. Assim, e dependendo do campo de aplicação, têm surgido muitos trabalhos com sugestões de amostradores. Gilks & Wild (1992) sugerem um método de rejeição adaptativo, podendo ser aplicado para amostrar de funções log-côncavas; Ritter & Tanner (1992) sugerem o método de inversão (*Griddy-Gibbs sampler*); Gilks *et al.* (1996) sugeriram o método híbrido - método de rejeição adaptativo com o passo de Metropolis, que é também aplicável a funções não log-côncavas; Neal (2003) desenvolveu algoritmos de amostragem em fatia (*slice sampler*) e Keith *et al.* (2008), usando métodos adaptativos, consideraram amostradores adaptativos independentes. É importante salientar que toda a literatura referida anteriormente tem por objetivo a implementação de métodos MCMC eficientes, em termos da velocidade de convergência, em modelos de natureza contínua, usualmente lineares e gaussianos, mas não deixando de lado os não-lineares e não-gaussianos.

A incorporação no modelo da incerteza associada às estimativas dos parâmetros

através da abordagem bayesiana, pode ser ainda expandida explorando alternativas para os algoritmos do método de Monte Carlo via Cadeias de Markov (MCMC), como por exemplo, o algoritmo *Reversible Jump*. Em relação às séries temporais, e mais especificamente, no caso de modelos autorregressivos, Brooks *et al.* (2003) e Vermaak *et al.* (2004) adaptaram o algoritmo de saltos reversíveis, RJMCMC, introduzido por Green (1995).

Este trabalho foi motivado pelo objetivo principal de implementar e desenvolver um método de Monte Carlo utilizando Cadeias de Markov apropriado no contexto das séries temporais de valores inteiros, especificamente para modelos de valores inteiros com limiares autoinduzidos introduzido por Monteiro (2010) e posteriormente publicado em Monteiro *et al.* (2012), seguido da construção de um modelo de valores inteiros com limiares autoinduzidos com estrutura periódica.

Para atingir os objetivos inicialmente traçados, numa primeira fase foi analisada a metodologia *Reversible Jump Markov Chain Monte Carlo* em modelos autorregressivos como apresentado em Brooks *et al.* (2003) e, já no âmbito de modelos de contagem, foi estudada a metodologia proposta por Neal & Subba Rao (2007) na qual se explorou a possibilidade de o amostrador depender de uma amostra com dados artificialmente aumentados (*Data Augmentation*).

As segunda e terceira fases do trabalho consistiram na construção de modelos autorregressivos de valores inteiros com limiares autoinduzidos, sem e com uma estrutura periódica: generalizando o modelo introduzido por Monteiro (2010) para ordens superiores em cada um dos regimes e considerando o modelo autorregressivo de valores inteiros com limiares autoinduzidos de ordem 1 com estrutura periódica.

Este trabalho está estruturado da seguinte forma:

O Capítulo 1 apresenta os fundamentos teóricos de: a) Modelos autorregressivos de valores inteiros; b) inferência bayesiana; c) métodos de Monte Carlo via Cadeias de Markov - MCMC d) Algoritmos de Ampliação de Dados. Faz-se uma breve síntese dos principais aspetos da teoria bayesiana e métodos MCMC, além das definições e estudos,

pertinentes a este trabalho, relacionados aos modelos autorregressivos de valores inteiros. O estudo bibliográfico permitiu um maior entendimento dos diferentes aspetos estudados anteriormente por diversos autores sobre os modelos **INAR** e métodos **MCMC**.

No Capítulo 2 apresenta-se o modelo autorregressivo de valores inteiros com limiares autoinduzidos, com ordens superiores à unidade nos diferentes ramos, faz-se o estudo das suas propriedades probabilísticas, a estimação de parâmetros - incluindo a seleção de modelos - do ponto de vista clássico ou bayesiano, seguido de um estudo de simulação e aplicação a dois conjuntos de dados reais.

O Capítulo 3 introduz o modelo autorregressivo de valores inteiros com limiares autoinduzidos com estrutura periódica; nele, além do estudo das propriedades probabilísticas do modelo, obtêm-se estimadores dos parâmetros (usando abordagem clássica) e analisam-se as suas propriedades assintóticas, determinam-se preditores pontuais h -passos à frente e complementa-se com um estudo de simulação.

Capítulo 1

Fundamentos Teóricos e Estado da Arte

“Today’s posterior distribution is tomorrow’s prior.”

Dennis Victor Lindley

O objetivo desse capítulo é descrever os principais conceitos relacionados à presente tese. São apresentadas as bases teóricas necessárias para capítulos posteriores. Parte-se da definição do operador *thinning* binomial e dos modelos autorregressivos de valores inteiros seguido de um breve apanhado sobre os principais aspetos da teoria bayesiana e dos métodos MCMC.

1.1 Operador *Thinning* Binomial

O operador *thinning* binomial, representado por “ \circ ”, proporciona uma operação a qual retorna sempre um número inteiro. Tal operador foi introduzido por Steutel & van Harn (1979) para adaptar o termo de variável aleatória auto-decomponível às séries de valores inteiros.

Definição 1.1. [*Operador thinning binomial*] Seja X uma variável aleatória discreta com suporte no conjunto \mathcal{N}_0 e $\alpha \in [0, 1]$. O operador *thinning* binomial entre α e X é dado por

$$\alpha \circ X := \sum_{i=1}^X U_i(\alpha),$$

onde $\{U_i(\alpha)\}$ é a sequência de variáveis aleatórias (v.a.) não negativas independentes e identicamente distribuídas (i.i.d.) de Bernoulli com probabilidade de sucesso α , e independente de X . A sucessão $\{U_i(\alpha)\}$ é chamada processo de contagem tal que

$$P_r(U_i = 1) = 1 - P_r(U_i = 0) = \alpha.$$

Resumidamente, $\alpha \circ X$ é o somatório de X v.as de Bernoulli i.i.d., cada uma das quais é unitária com probabilidade α . Assim $\alpha \circ X|X$ tem distribuição binomial com parâmetros (X, α) , onde X é o número de sucessos independentes e a probabilidade de cada sucesso é α .

Dada uma população com X elementos e α sendo a probabilidade de qualquer um dos elementos apresentar uma característica específica, o operador *thinning* binomial pode ser interpretado da seguinte forma: se um indivíduo desta população apresenta uma determinada característica de forma independente dos outros membros da população, então o número de elementos da população que apresentam esta característica é determinada por $\alpha \circ X$. Segundo Weiß (2008a), o operador *thinning* binomial é bastante atrativo não somente por ser matematicamente elegante mas também por ser de interpretação bastante intuitiva.

Este operador apresenta três propriedades que são obtidas, diretamente, da definição proposta por Steutel & van Harn (1979) as quais são apresentadas na Tabela 1.1.

Tabela 1.1: Propriedades do operador *thinning* binomial.

(i)	$0 \circ X = 0$	
(ii)	$1 \circ X = X$	
(iii)	$\alpha \circ (\beta \circ X) \stackrel{d}{=} (\alpha\beta) \circ X$	$\alpha, \beta \in [0, 1]$

As outras propriedades apresentadas na Tabela 1.2 foram estudadas por Oliveira

(2000), Silva & Oliveira (2004, 2005) e Silva (2005a) entre outros. Em todas as propriedades apresentadas X , Y e Z representam variáveis aleatórias independentes de valores inteiros não-negativos.

Tabela 1.2: Propriedades do operador *thinning* binomial, com $\alpha, \beta \in [0, 1]$

(iv)	$(\alpha + \beta) \circ X \stackrel{d}{\neq} (\alpha \circ X) + (\beta \circ X)$
(v)	$E[\alpha \circ X] = \alpha E[X]$
(vi)	$E[(\alpha \circ Y)X] = \alpha E[XY]$
(vii)	$E[(\alpha \circ X)YZ] = \alpha E[XYZ]$
(viii)	$E[(\alpha \circ X)^2] = \alpha^2 E[X^2] + \alpha(1 - \alpha)E[X]$
(ix)	$E[(\alpha \circ Y)^2X] = \alpha^2 E[XY^2] + \alpha(1 - \alpha)E[XY]$
(x)	$E[(\alpha \circ X)(\alpha \circ Y)] = \alpha\beta E[XY]$
(xi)	$E[XY(\beta \circ Z)] = \beta E[XYZ]$
(xii)	$E[X(\alpha \circ Y)(\beta \circ Z)] = \alpha\beta E[XYZ]$
(xiii)	$V[\alpha \circ X] = \alpha^2 V[X] + \alpha(1 - \alpha)E[X]$

A propriedade associativa da multiplicação usual também é válida para a operação *thinning* binomial (iii). No entanto a propriedade distributiva da multiplicação usual não pode ser aplicada à operação *thinning* binomial (iv). Note-se também que o operador *thinning* binomial introduz um termo acrescido na variância, dado por $\alpha(1 - \alpha)E[X]$. Este termo corresponde à variância de uma variável aleatória $Bi(E[X], \alpha)$. É também importante salientar que, em geral, os momentos de ordem superior a um que envolvem o operador *thinning* binomial também não são iguais aos respetivos momentos quando se usa a multiplicação usual em vez do referido operador.

Franke & Subba Rao (1995) estenderam o operador *thinning* binomial para o contexto multivariado.

Definição 1.2. [*Operador thinning binomial vetorial*] Seja $\mathbf{X} = [X_1, \dots, X_p]^T$ um vetor aleatório e \mathbf{A} uma matriz de dimensões $p \times p$, com elementos a_{ij} satisfazendo $0 \leq a_{ij} \leq 1$ para todo $i, j = 1, \dots, p$. O operador *thinning* binomial vetorial entre \mathbf{A} e \mathbf{X} é definida como um vetor aleatório cuja i -éssima componente é dada por

$$[\mathbf{A} \circ \mathbf{X}]_i := \sum_{j=1}^p a_{ij} \circ X_j, \quad i = 1, \dots, p,$$

onde todas as séries de contagem $a_{ij} \circ X_j$, são independentes.

As propriedades anteriores apresentadas para o caso univariado, são também satisfeitas para a operação *thinning* binomial vetorial e foram estudadas por Franke & Subba Rao (1995), Latour (1997) e Oliveira (2000). Algumas propriedades do operador *thinning* binomial vetorial, as quais são pertinentes no decorrer deste trabalho, são apresentadas na Tabela 1.3.

Tabela 1.3: Propriedades do operador *thinning* vetorial

(i)	$E[A \circ X] = AE[X]$
(ii)	$E[(A \circ X)(B \circ Y)^T] = AE[XY^T]B^T$
(iii)	$E[(A \circ X)(A \circ X)^T] = AE[XX^T]A^T + \text{diag}(CE[X])$

onde $\text{diag}(CE[X])$ representa a matriz diagonal $CE[X]$, onde $C = [c_{ij}]$, $c_{ij} = a_{ij}(1 - a_{ij})$.

1.2 Modelos Autorregressivos de Valores Inteiros (INAR)

O facto da operação *thinning* binomial retornar sempre um valor inteiro, proporcionou que McKenzie (1985) e Al-Osh & Alzaid (1987) propuseram um modelo autorregressivo de valores inteiros não negativos baseados nesta operação. A operação de multiplicação usual num modelo autorregressivo de primeira ordem, $\text{AR}(1)$, foi substituída pela operação *thinning* binomial, resultando num processo particular chamado de **INAR** (*INteger-valued AutoRegressive*). Os modelos **INAR** foram também extendidos para os casos de ordem $p > 1$ por Alzaid & Al-Osh (1990) e Du & Li (1991). Nas secções que se seguem os modelos $\text{INAR}(1)$ e $\text{INAR}(p)$ são apresentados.

1.2.1 Modelo $\text{INAR}(1)$

Um processo de contagem $\text{INAR}(1)$ define-se da seguinte maneira.

Definição 1.3. [*Modelo INAR(1)*] Seja (X_t) um processo estocástico de valor inteiro não negativo. Diz-se que (X_t) é um processo *INAR(1)* se X_t satisfaz a equação recursiva

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad (1.1)$$

com $\alpha \in [0, 1)$ e (ϵ_t) é uma sequência de variáveis aleatórias inteiras i.i.d.

A interpretação deste modelo é bastante intuitiva: o número de elementos no sistema no instante t , X_t , é a soma do número de elementos que permaneceram no sistema, cada um com probabilidade de sobrevivência α constante ao longo do tempo, até ao instante t como o número de elementos que entram no sistema durante o intervalo de tempo $(t-1, t]$.

É importante referir que o processo (X_t) em (1.1) é (estritamente) estacionário e algumas das propriedades do modelo como esperança, variância, funções de autocovariância e autocorrelação foram estudadas por Al-Osh & Alzaid (1987), Al-Osh & Alzaid (1988), McKenzie (1988) e Brännäs & Hellström (2001). Na Tabela 1.4 são apresentadas algumas propriedades importantes do modelo *INAR(1)*.

Tabela 1.4: Propriedades do modelo *INAR(1)*

(i)	$E[X_t] = \mu_\epsilon / (1 - \alpha), \mu_\epsilon = E[\epsilon_t]$
(ii)	$E[X_t X_{t-1}] = \alpha X_{t-1} + \mu_\epsilon$
(iii)	$V[X_t] = (\alpha \mu_\epsilon + \sigma_\epsilon^2) / (1 - \alpha^2), \sigma_\epsilon^2 = V[\epsilon_t]$
(iv)	$V[X_t X_{t-1}] = \alpha(1 - \alpha)X_{t-1} + \sigma_\epsilon^2$
(v)	$\gamma(h) := Cov(X_t, X_{t+h}) = \alpha^h V[X_t]$
(vi)	$\rho(h) := Corr(X_t, X_{t+h}) = \alpha^h$
(vii)	$P_{X_t}(s) = P_{\epsilon_t}(s)P_{X_t}(1 - \alpha + \alpha s)$, em que $P_X(s) := E[s^X]$

Como se pode observar na Tabela 1.4, o processo *INAR(1)* partilha algumas propriedades com o modelo *AR(1)*. Por exemplo, a função de autocorrelação (*FAC*) tende exponencialmente para zero quando $h \rightarrow \infty$. No entanto para o processo *INAR(1)* os valores da *FAC* são sempre positivos. Notar, porém, que a variância condicional $V[X_t | X_{t-1}]$ é uma função linear em X_{t-1} , situação esta, que não se verifica no caso convencional. Por outro lado, a partir da propriedade (vii) e do facto do processo

(X_t) ser estacionário, a distribuição marginal de (X_t) pode ser determinada a partir da relação

$$P_X(s) = P_\epsilon(s)P_X(1 - \alpha + \alpha s). \quad (1.2)$$

A relação anterior é importante uma vez que permite concluir que qualquer distribuição auto-decomponível para inteiros pode ser usada como distribuição marginal para X_t .

Por outro lado, da relação (1.2) pode-se também concluir que a distribuição marginal de (X_t) é completamente especificada a partir da distribuição de ϵ_t . Uma questão que habitualmente se coloca em relação à distribuição marginal de (X_t) é saber em que casos as distribuições de ϵ_t e X_t pertencem à mesma família. A este respeito, Al-Osh & Alzaid (1988) mostraram que as duas distribuições são do mesmo tipo se e só se uma delas pertencer à classe de distribuições discretas estáveis. No caso particular em que se quer garantir que $E[X_t] < \infty$, Al-Osh & Alzaid (1987) mostraram que para o modelo $\text{INAR}(1)$, a distribuição Poisson desempenha um papel análogo ao da distribuição Normal nos modelos ARMA de valores contínuos, isto é, a distribuição de X_t será $\text{Po}(\mu)$, com $\mu := \lambda/(1 - \alpha)$, se e só se ϵ_t tiver distribuição $\text{Po}(\lambda)$.

1.2.2 Modelo $\text{INAR}(p)$

De forma análoga aos modelos autorregressivos convencionais, no caso dos modelos INAR também é possível considerar modelos com ordens superiores a um. Alzaid & Al-Osh (1990) e Du & Li (1991) generalizaram o modelo $\text{INAR}(1)$ definido em (1.1) a uma ordem $p > 1$ e obtiveram o modelo $\text{INAR}(p)$.

Definição 1.4. [*Modelo $\text{INAR}(p)$*] Seja (X_t) um processo estocástico de valor inteiro não negativo. X_t é chamado de processo $\text{INAR}(p)$ se satisfaz a seguinte condição

$$X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \cdots + \alpha_p \circ X_{t-p} + \epsilon_t, \quad (1.3)$$

com $\alpha_i \in [0, 1]$, $i = 1, \dots, p-1$, $\alpha_p \in (0, 1]$ e sendo (ϵ_t) uma sucessão de variáveis aleatórias i.i.d. de valores inteiros não negativos com média e variância finitas, onde para cada instante t , ϵ_t é independente de X_{t-i} .

Cada um dos trabalhos acima citados apresentam a generalização do modelo $\text{INAR}(p)$ sob diferentes interpretações. No caso do modelo $\text{INAR}(p)$ proposto por Alzaid & Al-Osh (1990), os autores supuseram que a soma dos coeficientes associados aos operadores *thinning* binomial é inferior a um, isto é $\sum_{i=1}^p \alpha_i < 1$, e que cada uma das séries de contagem, dado X_t , é independente da história passada do processo e dos elementos que permanecem no sistema. Além disso, estes autores também assumiram que a distribuição condicionada do vetor $(\alpha_1 \circ_{t+1} X_t, \alpha_2 \circ_{t+2} X_t, \dots, \alpha_p \circ_{t+p} X_t)$, dado X_t é multinomial de parâmetros $(X_t, \alpha_1, \alpha_2, \dots, \alpha_p)$. Sob esta formulação, o modelo $\text{INAR}(p)$ tem a mesma estrutura de segunda ordem que um modelo $\text{ARMA}(p, p-1)$.

A perspectiva adotada por Du & Li (1991) admite que o modelo (1.3) pode ser tratado como um processo de renascimento múltiplo com migração, num período t fixo, onde todas as séries de contagem $\alpha_i \circ X_{t-i}$, $i = 1, \dots, p$, são mutuamente independentes e independente de ϵ_t . Du & Li (1991) mostraram que as condições de estacionaridade (de segunda ordem) do modelo $\text{INAR}(p)$ assim definido são as mesmas do modelo $\text{AR}(p)$, isto é, o processo é estacionário quando as raízes do polinómio característico de ordem p

$$\varphi[\lambda] = \lambda^p - \alpha_1 \lambda^{p-1} - \dots - \alpha_{p-1} \lambda - \alpha_p = 0, \quad (1.4)$$

estão dentro do círculo unitário e verificaram ainda que o processo $\text{INAR}(p)$ é ergódico. Dion *et al.* (1995) provaram que a condição $\sum_{i=1}^p \alpha_i < 1$, implica que todas as raízes do polinómio característico $\varphi[\lambda]$ estão dentro do círculo unitário.

Como representação alternativa, Franke & Subba Rao (1995) apresentaram o processo $\text{INAR}(p)$ como sendo um processo $\text{INAR}(1)$ p -dimensional.

Latour (1997) descreve um processo $\text{INAR}(p)$ m -variado, o $\text{MGINAR}(p)$, na forma

vetorial. Associando o processo $\text{INAR}(p)$ à definição do processo $\text{MGINAR}(p)$ dada por Latour (1997), pode-se dizer que o processo $\text{INAR}(p)$ pode ser visto como um subvetor de um processo de dimensão p multivariado $\text{INAR}(1)$.

Definição 1.5. [*Modelo $\text{INAR}(p)$ - representação vetorial*] Fazendo uso da definição de operador *thinning* dado em (1.2) pode-se reescrever (1.4) na forma vetorial

$$\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-1} + \mathbf{Z}_t, \quad (1.5)$$

onde $\mathbf{X}_t = [X_t \ X_{t-1} \ X_{t-2} \ \cdots \ X_{t-p+1}]^T$, $\mathbf{Z}_t = [\epsilon_t \ 0 \ \cdots \ 0]^T$, $a_i \in [0, 1]$, $i = 1, \dots, p$, $a_p \geq 0$ e

$$\mathbf{A} = \begin{pmatrix} a_1 & \cdots & a_{p-1} & a_p \\ 1 & \cdots & 0 & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}.$$

A partir desta representação do processo $\text{INAR}(p)$, Franke & Subba Rao (1995) demonstram a existência de uma solução estacionária para o processo.

Teorema 1.1. [*Franke & Subba Rao, 1995*] Seja (\mathbf{X}_t) , que satisfaz (2.3), uma cadeia de Markov irredutível e aperiódica em \mathbb{N}_0^p . Se $E\|\mathbf{Z}_t\| < +\infty$ e se o maior vetor próprio ρ de \mathbf{A} é inferior a um, então existe uma solução estritamente estacionária que satisfaz (1.5).

Franke & Subba Rao (1995) provaram também que qualquer solução de (\mathbf{X}_t) que satisfaz (1.5) é uma cadeia de Markov irredutível e aperiódica.

Lema 1.1. [*Franke & Subba Rao, 1995*] Para $0 < a_i < 1$, $i = 1, \dots, p$ e $0 < q^\beta(0) := P(\epsilon_t = 0) < 1$, então qualquer solução (\mathbf{X}_t) de (1.5) é uma cadeia de Markov irredutível e aperiódica em \mathbb{N}_0^p .

1.2.3 Estado da Arte

De acordo com o que foi mencionado anteriormente, os modelos de séries temporais para dados de contagem foram estudados por diversos autores. McKenzie (1986), Al-Osh & Alzaid (1987), Du & Li (1991) e Franke & Subba Rao (1995) introduziram os modelos INAR univariados e multivariados e estudaram algumas de suas propriedades.

Al-Osh & Alzaid (1987) estudaram os modelos PoInar(1) e propuseram estimadores de Yule-Walker, máxima verossimilhança e mínimos quadrados condicionais para estimar α e λ . Em Franke & Seligmann (1993) é possível verificar que o processo é estritamente estacionário e em Brännäs (1994, 1995) foi introduzido o método dos momentos generalizados. Posteriormente Freeland (1998) tratou de uma forma unificada o método dos mínimos quadrados condicional ponderado e o método da máxima verossimilhança condicional. Silva (2005b) fez a comparação entre estimativas clássicas e bayesianas em simulação e para conjuntos de dados reais usados em Freeland (1998).

McKenzie (1986) propôs o uso de modelos com distribuição geométrica e binomial negativa por analogia com os modelos AR com distribuição exponencial negativa e gama introduzindo os modelos INMA(1) (*INteger Moving Average*). O trabalho de McKenzie (1986) também considera o modelo INAR(1) com coeficientes não constantes e INAR(1) de margens com distribuição binomial negativa e usa a suposição que os coeficientes apresentam distribuição beta.

Al-Osh & Alzaid (1991) sugerem o uso de uma distribuição binomial como distribuição marginal do processo. Alzaid & Al-Osh (1993) apresentam uma forma alternativa para o modelo INAR(1) baseada numa operação *thinning* hipergeométrica.

Em Al-Osh & Aly (1992) também são estudadas as propriedades dos modelos INAR(1). Brännäs & Hellström (2001) atenuaram algumas condições de independência aplicadas a dados económicos e sugerem o uso de modelos INARCH e os modelos INAR(1) por limiares. Franke & Seligmann (1993) introduziram os modelos SINAR (*Switching INAR*) em que as inovações ϵ_t correspondem a dois estados representados

por variáveis de Poisson com intensidades λ_1 e λ_2 .

Latour (1998) generaliza o modelo $\text{INAR}(p)$ como $\text{GINAR}(p)$ ao substituir a variável aleatória de Bernoulli que consta na operação *thinning* por uma variável aleatória não negativa qualquer com momento de segunda ordem finito. Outros autores ainda demonstraram a relação do modelo $\text{GINAR}(p)$ a processos de ramificação multiplicativos com imigração e também indicam uma extensão a um $\text{GINARMA}(p, q)$.

No que diz respeito aos modelos $\text{INAR}(p)$, algumas extensões foram estudadas por Brännäs (1995), fazendo uso de variáveis explicativas, Berglund & Brännäs (1996) trabalhando com dados em painel e Silva *et al.* (2005) para dados replicados. Franke & Subba Rao (1995) generalizaram o modelo $\text{INAR}(1)$ para o caso multivariado, discutiram condições de estacionariedade e estudaram as propriedades dos estimadores dos mínimos quadrados condicionais para os parâmetros do modelo.

Para modelos auto-regressivos de ordem um de valores inteiros baseados no operador *thinning* cujos coeficientes variam de uma forma aleatória ao longo do tempo, considera-se como formulação inicial o modelo de Brännäs (1995) com duas modificações: a primeira é substituir o *thinning* binomial por uma generalização - (operador *thinning* generalizado, Gomes & Canto e Castro, 2009) e a segunda é em lugar de usar um vetor de covariáveis fixas, considerar uma sucessão de variáveis aleatórias independentes e identicamente distribuídas.

No domínio da frequência, os estimadores de Whittle foram propostos por Silva & Oliveira (2004). As estimativas bayesianas foram introduzidas por Silva & Silva (2004) e Silva *et al.* (2005). Outras distribuições foram consideradas na literatura, além da de Poisson, como a binomial, a binomial negativa, a geométrica ou a Poisson generalizada (McKenzie, 1986; Alzaid & Al-Osh, 1993; Brännäs, 1994).

Dos resultados que se conhecem da análise bayesiana aplicados a séries temporais discretas para modelos paramétricos, podem-se destacar West & Migon (1995), Chen & Ibrahim (2000), Durbin & Koopman (2000) e Man-Suk & Bin Lim (2001). A análise bayesiana dos modelos auto-regressivos de valores inteiros ainda é limitada.

Um deles é o de McCabe & Martin (2005) onde são feitas previsões para observações futuras do modelo INAR(1) sob o ponto de vista bayesiano considerando distribuições *a priori* uniforme para todos os parâmetros. Neal & Subba Rao (2007) apresentam um algoritmo MCMC considerado eficiente para as estimativas dos coeficientes de um processo INAR(p), sendo em Enciso-Mora *et al.* (2008) aplicado tal algoritmo para estimar as ordens p e q num modelo INARMA. Brännäs (1995) e Enciso-Mora *et al.* (2009) incorporam variáveis explicativas no seu estudo. Monteiro *et al.* (2008) aplicaram a metodologia clássica e bayesiana na obtenção de sistemas de alarme ótimo para processos de contagem INAR(1). O número reduzido de trabalhos bayesianos para a classe dos modelos INAR deve-se principalmente ao facto da função de verosimilhança ser, em geral, complexa.

1.3 Elementos Fundamentais de Estatística Bayesiana

Um dos objetivos deste trabalho é através de uma abordagem bayesiana obter estimativas dos parâmetros do modelo em estudo. Para isso torna-se necessário desenvolver e implementar um método de Monte Carlo utilizando Cadeias de Markov apropriado no contexto das séries temporais de valores discretos, em específico para modelos de valores inteiros com limiares autoinduzidos introduzidos por Monteiro *et al.* (2008).

Apesar de a estatística bayesiana estar bem documentada na literatura (ver, por exemplo, Paulino *et al.*, 2003), por uma questão de clarificação do trabalho, apresentam-se a seguir alguns conceitos fundamentais da inferência bayesiana, utilizados neste trabalho.

Na inferência bayesiana o parâmetro θ é encarado como um dos possíveis valores de um vetor aleatório. Assim, toda a informação *a priori* é quantificada em termos de uma distribuição de probabilidades $\pi(\theta)$ chamada distribuição *a priori* (Jeffreys,

1998). Esta distribuição *a priori* pode ser considerada como uma medida de credibilidade que é atribuída pelo investigador na análise de θ , baseada no conhecimento disponível sobre o fenómeno em estudo.

Além da informação *a priori*, considera-se também a informação obtida através da amostra, quantificada por $f(x|\theta)$. Combinando a distribuição e a função de verosimilhança através do Teorema de Bayes, é possível atualizar a informação sobre o parâmetro θ , mediante o conhecimento dos dados. Esta distribuição de probabilidade é denominada de distribuição *a posteriori* para θ . Assim, a partir de uma amostra aleatória $x := (x_1, x_2, \dots, x_n)$, de um modelo com função densidade de probabilidade $f(x|\theta)$, obtem-se a distribuição *a posteriori* através do Teorema de Bayes:

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)},$$

onde o denominador $f(x)$ representa a distribuição marginal de x

$$\pi(x) = \int f(x, \theta) d\theta = \int f(x|\theta)\pi(\theta) d\theta. \quad (1.6)$$

Quando não for necessária a especificação completa da distribuição *a posteriori*, usa-se a densidade *a posteriori* não normalizada

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta),$$

evitando-se o cálculo do integral (1.6), que por vezes só é possível recorrendo a métodos de integração numérica ou de simulação estocástica (Gamerman, 1996). Os métodos de simulação de Monte Carlo via Cadeias de Markov -MCMC (*Markov Chain Monte Carlo*) têm-se popularizado em inferência bayesiana como uma alternativa apropriada e eficiente para a resolução de integrais principalmente em ambientes multidimensionais (Paulino *et al.*, 2003; Gelman e Rubin, 1992).

1.4 Métodos de Monte Carlo via Cadeias de Markov

Apesar de o método de Monte Carlo (MC) ter uma aplicação mais alargada, vai-se ser focalizado no âmbito da metodologia bayesiana. Uma vez que na metodologia bayesiana a distribuição *a posteriori* é a base de toda a inferência, nomeadamente a obtenção das próprias estimativas dos parâmetros baseia-se no cálculo de medidas de localização, torna-se imprescindível calcular momentos e de um modo particular da distribuição *a posteriori*. Neste contexto é fundamental poderem calcular-se integrais do tipo

$$E[g(\theta)|x] = \int \pi(\theta|x)g(\theta)d\theta,$$

ou suas aproximações.

O método de MC permite obter um valor aproximado para o cálculo deste integral, calculando a média amostral da quantidade de interesse através da geração de uma amostra aleatória $(\theta_1, \dots, \theta_n)$ da distribuição *a posteriori* $\pi(\theta|x)$. Assim

$$E[g(\theta)|x] = \int \pi(\theta|x)g(\theta)d\theta \approx \frac{1}{n} \sum_{t=1}^n g(\theta_t),$$

uma vez que,

$$\frac{1}{n} \sum_{t=1}^n g(\theta_t) \xrightarrow{q.c} E[g(\theta)|x].$$

A limitação do método de MC reside na dificuldade de gerar amostras da distribuição de interesse $\pi(\theta|x)$ quando a sua expressão analítica é muito complexa ou θ é um vetor de dimensão elevada (Paulino *et al.*, 2003).

Em conjunto com os métodos de MC, o conceito de *Cadeia de Markov* é importante por estar inserido na teoria de processos estocásticos que leva em consideração a interdependência temporal das variáveis aleatórias e o seu comportamento limite (Gamerman, 1996).

Os métodos de MC via cadeias de Markov, referidos simplesmente como MCMC

(*Markov Chain Monte Carlo*) constituem um dos mais notáveis avanços na estatística nos últimos 30 anos. As suas origens remontam aos trabalhos de Metropolis *et al.* (1953) e Hastings (1970). No entanto, foi apenas nos finais dos anos oitenta que esta ideia começou a ser alvo de investigação. Inicialmente os métodos MCMC foram usados na estatística bayesiana, no entanto a sua utilização é mais alargada, podendo ser aplicada na dita estatística clássica. De uma forma geral aplicam-se os algoritmos MCMC para obter amostras da distribuição *a posteriori* dos parâmetros do modelo.

Assim, o objetivo do método MCMC consiste em simular valores de uma cadeia de Markov que, sob condições de regularidade, terá como distribuição estacionária a distribuição de $\pi(\theta|x)$ da qual se pretende obter uma amostra aleatória. Os métodos MCMC requerem que a cadeia satisfaça algumas características, nomeadamente:

- ser *homogénea* no tempo, ou seja, as probabilidades de transição não dependem dos instantes de transição mas apenas do tempo decorrido;
- ser *irredutível* é possuir apenas uma classe de equivalência, ou seja, cada estado pode ser atingido a partir de qualquer outro num número finito de iterações;
- ser *aperiódica*, ou seja, não ter estados absorventes.

A metodologia de MCMC simplifica este cálculo fatorizando esta distribuição num conjunto de distribuições condicionais, de dimensão inferior e que podem ser mais facilmente simuladas. Como construir a matriz de transição e por quanto tempo a cadeia precisa ser percorrida até que essa amostra reflita a distribuição estacionária, são dois problemas relacionados e relevantes no processo. Assim, os métodos MCMC são úteis na construção da cadeia de Markov cuja distribuição estacionária é a distribuição $\pi(\theta|x)$.

Os algoritmos de Metropolis-Hastings (Metropolis *et al.*, 1953; Hastings, 1970) e Amostrador de Gibbs (Gelfand & Smith, 1990) são algumas alternativas mais utilizadas para a construção das cadeias apropriadas. Nestes algoritmos a probabilidade

de aceitação é obtida a partir da comparação de distribuições e assim só é possível para espaços paramétricos de mesma dimensão.

Em alguns casos pode-se admitir que o parâmetro θ pode variar ao longo dos modelos sendo necessária uma cadeia com espaço de estados que muda de dimensão no decorrer das iterações. Green (1995) sugeriu um algoritmo que pode ser visto como uma generalização dos métodos MCMC onde são possíveis movimentos entre espaços paramétricos de diferentes dimensões.

MCMC com Saltos Reversíveis (RJMCMC)

O algoritmo RJMCMC (*Reversible Jump Markov Chain Monte Carlo*) permite efetuar *saltos* entre os espaços e ainda manter a equação de equilíbrio detalhada de maneira suficiente para garantir a irreversibilidade da cadeia. O algoritmo RJMCMC move-se entre espaços de parâmetros de um conjunto de modelos, em geral em número finito e na maioria das situações são modelos encaixados. Assim o método é uma combinação do algoritmo MCMC para um dado modelo (fixo) com o passo adicional que permite a movimentação entre diferentes modelos. Por uma questão de simplificação considera-se a situação de os modelos estarem encaixados, que é a situação mais frequentemente usada. Isto significa incluir ou retirar parâmetros do modelo atual. Suponham-se K modelos possíveis, θ_k representa o conjunto de parâmetros do k -ésimo modelo, $k = 1, \dots, K$, i.e., de M_k .

- Algoritmo RJMCMC, com N iterações

1. Escolher um modelo inicial M_k , i.e., inicializar k e θ_k , na iteração $t = 1$.
2. Para cada iteração $t > 1$, fazer
 - (a) movimento dentro do modelo: para o modelo fixo M_k - obtido no fim da iteração $t - 1$ - atualizar os parâmetros θ_k , de acordo com a metodologia MCMC;

- (b) movimento entre modelos: propor a mudança para um novo modelo $M_{k'}$, calcular a probabilidade de aceitação para o movimento e decidir pela aceitação ou rejeição do novo modelo.

3. Incrementar a iteração $t = t + 1$; se $t < N$ voltar a 2.

• Movimento entre modelos

Admitam-se K modelos possíveis e θ_k os parâmetros do modelo k , para $k = 1, \dots, K$. Suponha-se que se está no modelo M_k de parâmetro θ_k e que se pretende fazer a mudança para o modelo $M_{k'}$ de parâmetro $\theta'_{k'}$. Sejam \mathbf{u}_k as variáveis auxiliares associadas ao modelo k , $k = 1, \dots, K$, tais que para qualquer modelo k , (θ_k, \mathbf{u}) tem sempre o mesmo número de elementos, por exemplo L . Para $1 \leq k, k' \leq K$, seja $g_{k \rightarrow k'}$ uma função invertível tal que $(\theta'_{k'}, \mathbf{u}') = g_{k \rightarrow k'}(\theta_k, \mathbf{u})$ com $g_{k' \rightarrow k} = g_{k \rightarrow k'}^{-1}$ e $r_k(\mathbf{u})$ a proponente para as variáveis auxiliares \mathbf{u} . A probabilidade de aceitação do movimento do modelo k com parâmetro θ_k , para o modelo k' com parâmetro $\theta'_{k'}$, é

$$\alpha(\theta_k, \theta'_{k'}) := \min \left\{ 1, \frac{\pi(k', \theta'_{k'} | \mathbf{x})}{\pi(k, \theta_k | \mathbf{x})} \times \frac{q(k | k')}{q(k' | k)} \times \frac{r_{k'}(\mathbf{u}')}{r_k(\mathbf{u})} \times J \right\},$$

onde $\pi(k', \theta'_{k'} | \mathbf{x})$ é a distribuição *a posteriori* dos parâmetros do modelo $M_{k'}$, $q(k' | k)$ é a probabilidade do movimento $k \rightarrow k'$, $r_k(\mathbf{u})$ a distribuição proponente para as variáveis auxiliares \mathbf{u} e J o jacobiano da transformação do movimento descrito.

É necessário calcular o jacobiano da transformação sempre que se faz um movimento entre modelos com diferentes parâmetros. Este jacobiano é o determinante da matriz V de dimensão $L \times L$ definida por

$$V = \left[\frac{\partial g_{k \rightarrow k'}(\theta_k, \mathbf{u})}{\partial(\theta_k, \mathbf{u})} \right].$$

Evidentemente que o movimento inverso de $M_{k'}$ para M_k é feito deterministicamente e é aceite com probabilidade

$$\alpha(\theta'_{k'}, \theta_k) = \left(\frac{1}{\alpha(\theta_k, \theta'_{k'})} \right).$$

Admita-se que o novo modelo a propor $M_{k'}$ tem dimensão superior, i.e., $n_{k'} > n_k$. Seguem-se os seguintes passos:

- propor o movimento $M_k \rightarrow M_{k'}$ com probabilidade $q(k'|k)$;
- gerar vetor \mathbf{u} , $n_{k'} - n_k$ dimensional, a partir de $r_k(\mathbf{u})$;
- fazer $(\theta'_{k'}, \mathbf{u}') = g_{k \rightarrow k'}(\theta_k, \mathbf{u})$;
- aceitar $(\theta'_{k'}, \mathbf{u}')$ com probabilidade $\alpha(\theta_k, \theta'_{k'})$.

1.5 Algoritmos de Ampliação de Dados

O termo *ampliação de dados*, designado pela sigla DA - (*Data Augmentation*) faz referência aos métodos que visam a construção de otimizações iterativas ou algoritmos de amostragem, através da introdução de variáveis latentes ou de dados não observados (Dyk & Meng, 2001). Os algoritmos DA partilham uma abordagem comum: quando a verosimilhança ou a *a posteriori* são difíceis de lidar em lugar de se fazer uma maximização ou simulação mais complicada, é possível através destes algoritmos simplificar os cálculos envolvidos (Tanner, 1996). Aumentam-se os dados observados com variáveis latentes possibilitando que a verosimilhança ou a distribuição *a posteriori* se tornem tratáveis. Esses algoritmos requerem que sejam feitas, no mínimo, duas suposições: uma sobre a distribuição dos dados e outra sobre o mecanismo associado aos dados omissos.

Dois algoritmos bastante difundidos neste contexto são o *Expectation-Maximization* (EM) (Dempster *et al.*, 1977) e o *Data Augmentation* (Tanner & Wong, 1987). Quando o objetivo do estudo é maximizar a verosimilhança ou a distribuição *a posteriori* o

algoritmo apropriado é o EM. Se o objetivo é explorar a verosimilhança ou distribuição *a posteriori*, depois de se fazer a ampliação dos dados, a abordagem escolhida deve ser a do algoritmo DA.

O algoritmo EM maximiza a verosimilhança dos dados observados usando os dados completos $Y = \{X, Z\}$ de maneira conveniente e tornou-se bastante popular entre os algoritmos determinísticos. Entre os algoritmos estocásticos têm-se o algoritmo DA que explora a densidade *a posteriori* $p(\theta|x)$ usando a densidade condicional $p(\theta|x, z)$.

1.5.1 Algoritmo *Data Augmentation*

No contexto bayesiano os dados omissos e os parâmetros do modelo são tratados da mesma maneira. Da mesma forma que se faz a inferência *a posteriori* dos parâmetros do modelo, pode-se também fazer a inferência *a posteriori* sobre as variáveis não observadas.

O algoritmo DA apresenta como ideia principal determinar a distribuição *a posteriori* do parâmetro θ , condicional aos dados observados X , $\pi(\theta|x)$, além de supor que se pode ampliar os dados X em $\{X, Z\}$ de tal modo que a distribuição *a posteriori* dos dados ampliados seja tratável. Tal algoritmo baseia-se na interação entre duas identidades: a identidade da *a posteriori* e a identidade preditiva.

A identidade da *a posteriori* é dada por

$$\pi(\theta|x) = \int \pi(\theta|x, z)p(z|x)dz,$$

onde $p(z|x)$ é a densidade preditiva dos dados latentes, dado x , e resulta da identidade preditiva dada por

$$p(z|x) = \int_{\theta} p(z|\theta, x)\pi(\theta|x)d\theta.$$

Para implementar este procedimento, Tanner & Wong (1987) consideraram a integração de MC, sendo o algoritmo definido da seguinte maneira:

- Simular uma amostra i.i.d. z_1, z_2, \dots, z_m de $p(z|x)$;

- Atualizar a aproximação atual de $P(\theta|x)$ pela média das *a posteriori* ampliadas em Z

$$\pi_{i+1}(\theta|x) \simeq \frac{1}{m} \sum_{j=1}^m p(\theta|x, z_i).$$

- A partir de $\pi_{i+1}(\theta|x)$, obter m novos θ_i^* ;
- Gerar o próximo z_i 's de $p(z|\theta_i^*, x)$. Usar estes z_i no primeiro passo e repetir o processo até obter a convergência.

1.5.2 Algoritmo Expectation-Maximization (EM)

O algoritmo **EM** (*Expectation Maximization*) é um método que permite calcular as estimativas de máxima verosimilhança dos parâmetros de um modelo, na presença de variáveis não observadas. Trata-se de um método de ampliação de dados; os dados observados são ampliados com dados latentes (i.e, não observados), com o objetivo de simplificar o problema de cálculo. É, pois, um algoritmo de grande utilidade quando os modelos dependem de dados incompletos.

Dempster *et al.* (1977) fizeram a descrição e análise formal do método **EM** ressaltando que, ao longo dos anos, este método havia sido proposto muitas vezes em circunstâncias especiais em diversos trabalhos. Detalhes adicionais e inúmeras alternativas para o **EM** podem ser encontradas em McLachlan & Krishnan (1997).

A maior parte das aplicações do algoritmo **EM** fazem-se pelo facto de a própria estrutura dos dados sugerir a existência de dados omissos, dados censurados ou de mistura de distribuições. Noutras situações a complexidade do cálculo para a obtenção de estimativas de máxima verosimilhança leva a que se procure alternativas, por exemplo através da introdução de variáveis latentes no processo e daí esta metodologia ser aplicada.

Seja $X = (x_1, \dots, x_n)$ o conjunto de dados observados e Z uma variável aleatória cuja distribuição de probabilidade depende dos parâmetros desconhecidos θ e dos

dados observados. Os dados completos $Y = (X, Z)$ são definidos como os dados observados X e os dados aumentados Z , os quais podem ser uma variável latente ou dados incompletos. Denota-se por $L(y; \theta) = p(x, z|\theta)$ a função de verosimilhança dos dados completos; a estimativa de máxima verosimilhança dos parâmetros desconhecidos é determinada através da verosimilhança marginal dos dados observados

$$L(x; \theta) = \pi(x|\theta) = \int p(x, z|\theta) dz.$$

De forma resumida o algoritmo **EM** é definido por dois passos: no passo "E" o algoritmo calcula o valor esperado do logaritmo da verosimilhança e no passo "M" encontra seu máximo.

Algoritmo EM

Entrada: uma estimativa inicial $\theta^{(0)}$

Saída: sequência de $\theta^{(t)}, t = 1, 2, \dots$ de estimativas iteradas para θ

- Passo E (*Expectation*)

$$\begin{aligned} Q(\theta|\theta^{(t)}) &= \int_z \ln[L(y; \theta)] p(z|x, \theta^{(t)}) dz \\ &= E_Z[\ln L(y; \theta)|x, \theta^{(t)}]. \end{aligned}$$

- Passo M (*Maximization*)

$$\theta^{(t+1)} = \mathbf{argmax} Q(\theta|\theta^{(t)}),$$

faça-se $t = t + 1$ até que $\theta^{(t+1)} = \theta^{(t)}$, para determinada precisão.

O critério de paragem pode ser definido como $\| \theta^{(t+1)} - \theta^{(t)} \| < \delta$, com $\delta > 0$ e suficientemente pequeno. A sequência $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots$, converge para a estima-

tiva de máxima verossimilhança global quando a função Q tem um único máximo, de outra forma converge para o máximo local. A taxa de convergência nas primeiras iterações é em geral bastante boa, mas pode tornar-se muito lenta se se aproximar de um máximo local. No trabalho de Dempster *et al.* (1977) são apresentados alguns resultados que demonstram que a taxa de convergência do algoritmo EM é linear além de depender da proporção de informação nos dados observado. Assim, quando comparado a um problema de dados completos, se existe uma grande proporção de dados incompletos a convergência pode ser bastante lenta. Wu (1983) investigou pormenorizadamente diversos pontos do algoritmo EM e afirma que sob condições bastante gerais, o algoritmo tem convergência global confiável, isto é, a convergência é quase sempre num máximo local.

Casella & Berger (2001) sugerem que o algoritmo EM converge quase certamente para o MLE. Quando existem múltiplos pontos de máximo, o algoritmo EM, como outros métodos iterativos, não garante a convergência para o máximo global além do facto, da estimativa obtida depender do valor inicial. A taxa de convergência é geralmente mais lenta quando comparada com outros métodos disponíveis.

Em resumo considera-se o algoritmo EM com grande utilidade por apresentar simplicidade conceptual e facilidade de implementação. Como desvantagem apresenta a lenta convergência quando os dados são incompletos e em algumas situações o passo E pode ser intratável analiticamente sendo necessária uma aproximação por métodos de MC.

1.5.3 Algoritmo *Monte Carlo Expectation Maximization*

O algoritmo *Monte Carlo Expectation Maximization* - MCEM (Wei & Tanner, 1990) é uma modificação do algoritmo EM no qual o valor esperado do passo E é calculado numericamente através de simulações de MC. Desta forma, pode-se dizer que o algoritmo MCEM é uma aproximação ao algoritmo EM onde o cálculo do valor esperado é substituído pela média de m elementos gerados aleatoriamente da distribuição de

probabilidade das variáveis latentes z dado os valores observados x .

O algoritmo MCEM é definido pelos seguintes passos:

- Simular uma amostra i.i.d z_1, z_2, \dots, z_m de $p(z|x, \theta^{(t)})$;
- Calcular

$$Q(\theta|\theta^{(t)}) := \frac{1}{m} \sum_{j=1}^m \ln[L(z_j; \theta)|x, \theta^{(t)}];$$

- Determinar

$$\theta^{(t+1)} = \mathbf{argmax} Q(\theta|\theta^{(t)}).$$

Nesta abordagem, a especificação do valor de m pode ser considerado um problema. Uma possível solução consiste em aumentar o valor de m à medida que o número de iterações cresce e ao mesmo tempo monitorizar a convergência como especificado anteriormente para o algoritmo EM.

1.5.4 Aplicações dos Algoritmo RJMCMC e EM em Séries de Valores Inteiros

No que diz respeito a séries autorregressivas de valores inteiros, Pavlopoulos & Karlis (2008) usaram um algoritmo EM apropriado para estimar numericamente as estimativas de máxima verosimilhança dos parâmetros θ de uma mistura finita de m variáveis aleatórias Poisson-Binomial. Dada uma estimativa inicial $\theta^{(0)}$, admissível no espaço do parâmetro, o algoritmo EM estima primeiro o valor esperado condicional da informação latente dados os valores observados e os valores atuais dos parâmetros θ . O passo M atualiza θ com os novos valores θ' através da máxima verosimilhança dos dados aumentados. No trabalho realizado por Pavlopoulos & Karlis (2008) foram encontradas expressões explícitas para todos os parâmetros do modelo. Enciso-Mora *et al.* (2008) utilizaram o algoritmo EM para estimar os parâmetros do modelo

$\text{INAR}(p)$, com p fixo, e através do critério de seleção de modelos BIC, compararam diferentes modelos $\text{INAR}(p)$. No mesmo artigo, os autores verificaram que para modelos $\text{INARMA}(p, q)$ o uso de um algoritmo EM, como apresentado até agora, é inviável devido a dependência de alguns valores no passo E sendo possível, no entanto, a utilização de um algoritmo EM de MC. O algoritmo EM de MC pode apresentar também um mau desempenho devido a convergência ser lenta e pelo facto do amostrador utilizado para as estimativas no passo E apresentar uma elevada taxa de rejeição. Os autores também constataram que o algoritmo EM para modelos $\text{INAR}(p)$ gerou resultados bastante similares aos outros algoritmos MCMC aplicados.

Com o objetivo de criarem um conjunto de dados completos Anderson & Karlis (2010) utilizaram o método bridge que permite gerar os dados em falta no modelo e em todos os dados disponíveis usando o condicionamento. Após isso usaram o algoritmo EM com passo de MC para encontrar as estimativas dos parâmetros desejados.

Tendo em consideração o algoritmo RJMCMC, Neal & Subba Rao (2007) apresentam um algoritmo considerado eficiente para obter estimativas dos coeficientes de um processo $\text{INAR}(p)$ e da ordem correspondente.

Capítulo 2

Processos Autorregressivos de Valores Inteiros Definidos por Limiares Autoinduzidos

*“The essence of mathematics is not to make
simple things complicated, but to make
complicated things simple.”*

Stanley Gudder

O objetivo deste capítulo é generalizar a classe de modelos $\text{SETINAR}(2; 1, 1)$ introduzida por Monteiro (2010) e Monteiro *et al.* (2012) para ordens superiores. Tal classe de modelos permite caracterizar séries de contagem que apresentam mudanças da volatilidade no tempo e padrões do tipo linear que se repetem por troços. O uso de limiares permite representar uma série temporal por meio de modelos autorregressivos de valores inteiros por partes dividindo o conjunto de dados em grupos que podem ser chamados regimes.

O modelo aqui estudado, designado por $\text{SETINAR}(2; p^{(1)}, p^{(2)})$, apresenta dois regimes cujas ordens das componentes autorregressivas de cada ramo são dadas por $p^{(1)}$ e $p^{(2)}$ e as inovações são variáveis aleatórias independentes com distribuição de Poisson.

Na Secção 2.1 é definido o modelo e apresentada a sua forma vetorial; na Secção 2.2 são apresentadas as suas propriedades probabilísticas; a Secção 2.3.1 apresenta os estimadores de mínimos quadrados condicionais; na Secção 2.3.2 tem-se a verosimilhança seguida pela Inferência bayesiana na Secção 2.3.3; o estudo de simulação é feito na Secção 2.4 e conclui-se o capítulo com aplicação a dados reais na Secção 2.5.

2.1 Definição do Modelo SETINAR

A classe dos modelos autorregressivos de valores inteiros com limiar autoinduzido, SETINAR, do inglês *Self-Exciting Threshold INteger-valued Autoregressive*, é definida por modelos da forma

$$X_t = \begin{cases} \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} + Z_t^{(1)}, & \text{se } X_{t-d} \leq r \\ \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} + Z_t^{(2)}, & \text{se } X_{t-d} > r. \end{cases} \quad (2.1)$$

O modelo apresenta dois regimes ($k = 1, 2$) e um valor de limiar r . Assumem-se diferentes ordens autorregressivas $p^{(1)}$ e $p^{(2)}$ para cada um dos regimes e um parâmetro de desfasamento igual a d . Os coeficientes autorregressivos são definidos por

$$\alpha^{(1)} := [\alpha_1^{(1)} \cdots \alpha_{p^{(1)}}^{(1)}]^T, \quad \alpha^{(2)} := [\alpha_1^{(2)} \cdots \alpha_{p^{(2)}}^{(2)}]^T, \quad \text{com} \quad \sum_{i=1}^{p^{(k)}} \alpha_i^{(k)} < 1, \quad k = 1, 2.$$

O termo $Z_t^{(k)}$, com $t \in \mathbb{Z}$, é uma sequência de variáveis aleatórias i.i.d. independentes de X_t e de $\alpha^{(k)} \circ X_{t-d}$, com distribuição de Poisson de parâmetro $\lambda^{(k)}$. Assim, de forma geral, pode-se definir o modelo SETINAR($2; p^{(1)}, p^{(2)}$) em (2.1) da seguinte maneira.

Definição 2.1. [Modelo SETINAR($2; p^{(1)}, p^{(2)}$)] O modelo SETINAR($2; p^{(1)}, p^{(2)}$) é definido pela equação recursiva

$$X_t = \left(Y_t^{(1)} + Z_t^{(1)} \right) I_{t-d}^{(1)} + \left(Y_t^{(2)} + Z_t^{(2)} \right) I_{t-d}^{(2)},$$

com $\alpha_i^{(k)} \in [0, 1)$, $i = 1, \dots, p^{(k)}$, $k = 1, 2$, onde $Y_t^{(k)} := \sum_{i=1}^{p^{(k)}} \alpha_i^{(k)} \circ X_{t-i}$, $Y_{t,i}^{(k)} := \alpha_i^{(k)} \circ X_{t-i} | X_{t-i} \sim \text{Bin}(X_{t-i}, \alpha_i^{(k)})$, Z_t é uma sequência de v.a. i.i.d. e

$$I_{t-d}^{(1)} := \begin{cases} 1 & \text{se } X_{t-d} \leq r \\ 0 & \text{se } X_{t-d} > r \end{cases}, \quad I_{t-d}^{(2)} = 1 - I_{t-d}^{(1)}. \quad (2.2)$$

Desta forma o processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ pode ser escrito como um processo $\text{INAR}(p^{(1)} I_{t-d}^{(1)} + p^{(2)} I_{t-d}^{(2)})$, onde cada regime é um processo $\text{INAR}(p^{(k)})$, isto é,

$$X_t = \begin{cases} \text{INAR}(p^{(1)}), & \text{se } X_{t-d} \leq r \\ \text{INAR}(p^{(2)}), & \text{se } X_{t-d} > r \end{cases}.$$

2.1.1 Representação Vetorial do Modelo SETINAR

O processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ pode ser escrito na representação p -variada de um processo univariado $\text{INAR}(1)$ (Franke & Subba Rao, 1995), tal como estudado por Du & Li (1991) e aqui apresentado na Definição 1.4. Assim, a partir do operador *thinning* binomial vetorial apresentado na Definição 1.2 pode-se expressar o processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ como um processo p -dimensional $\text{INAR}(1)$.

Definição 2.2. [Representação Vetorial $\text{SETINAR}(2; p^{(1)}, p^{(2)})$] Seja (X_t) um processo estocástico de valores inteiros não negativos. Assim, X_t admite a representação

$$\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-1} + \mathbf{Z}_t, \quad (2.3)$$

onde \mathbf{A} é uma matriz $p \times p$, com $p := p^{(1)} I_{t-d}^{(1)} + p^{(2)} I_{t-d}^{(2)}$,

$$\mathbf{A} = \begin{pmatrix} \sum_{k=1}^2 \alpha_1^{(k)} I_{t-d}^{(k)} & \cdots & \sum_{k=1}^2 \alpha_{p-1}^{(k)} I_{t-d}^{(k)} & \sum_{k=1}^2 \alpha_p^{(k)} I_{t-d}^{(k)} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}, \quad (2.4)$$

sendo $\mathbf{X}_t = [X_t X_{t-1} \cdots X_{t-p+1}]^T$, e $\mathbf{Z}_t = \left[\sum_{k=1}^2 Z_t^{(k)} I_{t-d}^{(k)} 0 \cdots 0 \right]^T$, onde $Z_t^{(1)}$ e $Z_t^{(2)}$ são sequências de variáveis aleatórias inteiras independentes, e com distribuições F_1 e F_2 , respetivamente.

Esta representação mostra que o processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ pode ser visto como um subvetor de um processo de dimensão p -multivariado $\text{INAR}(1)$.

Na Figura 2.1 apresentam-se duas realizações do processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$, com valores distintos ou iguais entre os regimes, respetivamente para os parâmetros das inovações. No gráfico (a), as realizações movem-se bastante entre os dois regimes; nesta situação os valores de $\lambda^{(1)}$ e $\lambda^{(2)}$ são iguais. No gráfico (b), onde $\lambda^{(1)}$ e $\lambda^{(2)}$ são diferentes, as mudanças de regime são menos frequentes.

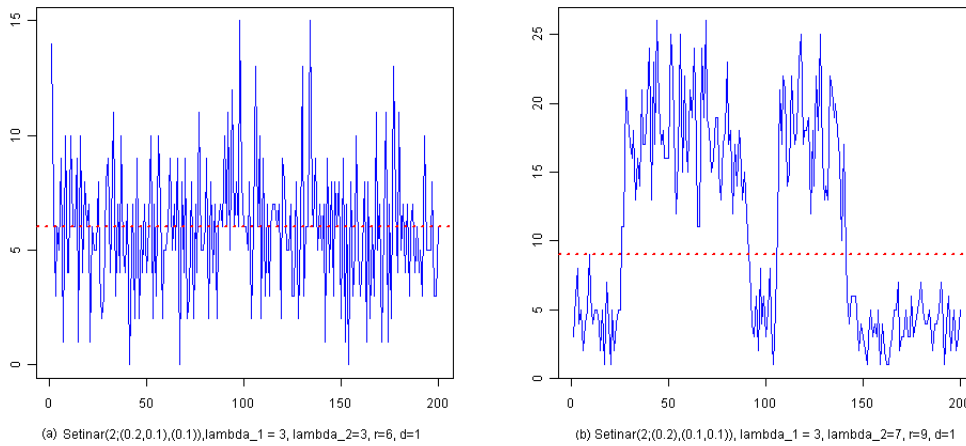


Figura 2.1: Realizações de dois modelos $\text{SETINAR}(2; p^{(1)}, p^{(2)})$.

2.2 Propriedades Gerais

2.2.1 Estacionaridade Estrita

Já se referiu anteriormente que o processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ pode-se expressar como um processo p -dimensional $\text{INAR}(1)$. Desta forma, para verificar-se a existência de um processo SETINAR estritamente estacionário que satisfaz (2.3), aplicam-se os resultados obtidos do Teorema 1 de Franke & Subba Rao (1995) para um processo $\text{INAR}(p)$.

Teorema 2.1 (Distribuição estritamente estacionária). *Seja (\mathbf{X}_t) , o processo definido em (2.3), uma cadeia de Markov irreduzível e aperiódica em \mathbb{N}_0^p . Se $E\|Z_t\| < +\infty$ e se o maior valor próprio ρ de \mathbf{A} é inferior a 1, então existe uma solução estritamente estacionária que satisfaz (2.3).*

Demonstração. Pela proposição B em Dion *et al.* (1995, p. 126) tem-se que

$$\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} I_{t-d}^{(1)} + \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} I_{t-d}^{(2)} < 1 \Leftrightarrow \rho < 1. \quad (2.5)$$

A condição (2.5) implica que todas as raízes do polinómio característico \mathbf{A} estão dentro do círculo unitário. Além disso, se $E\|Z_t\| < +\infty$, então pelo Teorema 1 em Franke & Subba Rao (1995) existe uma solução estritamente estacionária que satisfaz (2.3). ■

2.2.2 Propriedades Probabilísticas

Para determinar a média, variância e a função de autocovariância do processo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ define-se, $q := P(X_{t-d} \leq r)$, $u_{i,d}^{(1)} := E[X_{t-i} | X_{t-d} \leq r]$, $u_{i,d}^{(2)} := E[X_{t-i} | X_{t-d} > r]$, $\gamma_{s,h}^{(1)} := \text{Cov}[X_t, X_{t+s} | X_{t+h} \leq r]$, $\gamma_{s,h}^{(2)} := \text{Cov}[X_t, X_{t+s} | X_{t+h} > r]$.

Lema 2.1. Para o modelo $SETINAR(2; p^{(1)}, p^{(2)})$ a média é dada por

$$\mu := E[X_t] = q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i,d}^{(1)} + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i,d}^{(2)} + q\mu_Z^{(1)} + (1-q)\mu_Z^{(2)},$$

sendo $\mu_Z^{(k)} := E[Z_t^{(k)}]$ for $k = 1, 2$. A variância é

$$\begin{aligned} \sigma^2 := V[X_t] &= q \left\{ V \left[\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} X_{t-i} | X_{t-d} \leq r \right] + \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} (1 - \alpha_i^{(1)}) u_{i,d}^{(1)} \right\} + q\sigma_Z^{(1)} \\ &+ (1-q) \left\{ V \left[\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} X_{t-i} | X_{t-d} > r \right] + \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} (1 - \alpha_i^{(2)}) u_{i,d}^{(2)} \right\} + (1-q)\sigma_Z^{(2)} \\ &+ q(1-q) \left[\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i,d}^{(1)} - \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i,d}^{(2)} \right]^2, \end{aligned}$$

com $\sigma_Z^{(k)} := V[Z_t^{(k)}]$, para $k = 1, 2$. Finalmente, a função de autocovariância $\gamma(s)$ é dada por

$$\begin{aligned} \gamma(s) &= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \gamma_{s-i, s-d}^{(1)} + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \gamma_{s-i, s-d}^{(2)} + \\ &+ q(1-q)(u_{0,d-s}^{(1)} - u_{0,d-s}^{(2)}) \left\{ \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i-s, d-s}^{(1)} - \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i-s, d-s}^{(2)} \right\}. \end{aligned}$$

Demonstração. A média é dada por

$$\begin{aligned} E[X_t] &= E \left[\sum_{k=1}^2 \left(\sum_{i=1}^{p^{(k)}} \alpha_i^{(k)} \circ X_{t-i} + Z_t^{(k)} \right) I_{t-d}^{(k)} \right] \\ &= q \left[\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} E[X_{t-i} | X_{t-d} \leq r] + E[Z_t^{(1)}] \right] + \\ &+ (1-q) \left[\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} E[X_{t-i} | X_{t-d} > r] + E[Z_t^{(2)}] \right] \\ &= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i,d}^{(1)} + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i,d}^{(2)} + q\mu_Z^{(1)} + (1-q)\mu_Z^{(2)}. \end{aligned}$$

A variância é dada por

$$\begin{aligned}
 V[X_t] &= V\left[\sum_{k=1}^2\left(\sum_{i=1}^{p^{(k)}}\alpha_i^{(k)}\circ X_{t-i}\right)I_{t-d}^{(k)}\right] + V\left[\sum_{k=1}^2Z_t^{(k)}I_{t-d}^{(k)}\right] \\
 &= V\left[\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}\right)I_{t-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}}\alpha_i^{(2)}\circ X_{t-i}\right)I_{t-d}^{(2)}\right] + V\left[\sum_{k=1}^2Z_t^{(k)}I_{t-d}^{(k)}\right] \\
 &= V\left[\underbrace{\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}\right)I_{t-d}^{(1)}}_I + V\left(\underbrace{\sum_{i=1}^{p^{(2)}}\alpha_i^{(2)}\circ X_{t-i}}_{II}\right)I_{t-d}^{(2)}\right] + \\
 &\quad + 2Cov\left\{\underbrace{\left[\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}\right]I_{t-d}^{(1)}, \left[\sum_{i=1}^{p^{(2)}}\alpha_i^{(2)}\circ X_{t-i}\right]I_{t-d}^{(2)}}_{III}\right\} + V\left[\underbrace{\sum_{k=1}^2Z_t^{(k)}I_{t-d}^{(k)}}_{IV}\right].
 \end{aligned} \tag{2.6}$$

O termo I da variância é dado por

$$\begin{aligned}
 I &= V\left[\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}\right)I_{t-d}^{(1)}\right] \\
 &= V\left[E\left[\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}\right)I_{t-d}^{(1)}|X_{t-d}\right]\right] + E\left[V\left[\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}\right)I_{t-d}^{(1)}|X_{t-d}\right]\right] \\
 &= V\left[I_{t-d}^{(1)}E\left[\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}|X_{t-d}\right]\right] + E\left[I_{t-d}^{(1)}V\left[\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\circ X_{t-i}|X_{t-d}\right]\right] \\
 &= E\left[I_{t-d}^{(1)}\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}X_{t-i}\right)^2\right] - E\left[I_{t-d}^{(1)}\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}X_{t-i}\right]^2 + E\left[I_{t-d}^{(1)}\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}(1-\alpha_i^{(1)})X_{t-i}\right] \\
 &= P(X_{t-d} \leq r)E\left[\left(\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}X_{t-i}\right)^2|X_{t-d} \leq r\right] - \left[P(X_{t-d} \leq r)\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}\times\right. \\
 &\quad \left.\times E[X_{t-i}|X_{t-d} \leq r]\right]^2 + P(X_{t-d} \leq r)\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}(1-\alpha_i^{(1)})E[X_{t-i}|X_{t-d} \leq r] \\
 &= q\left\{V\left[\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}X_{t-i}|X_{t-d} \leq r\right] + \sum_{i=1}^{p^{(1)}}\alpha_i^{(1)}(1-\alpha_i^{(1)})u_{i,d}^{(1)}\right\} + q(1-q)\sum_{i=1}^{p^{(1)}}\alpha_i^{(1)2}u_{i,d}^{(1)2}.
 \end{aligned} \tag{2.7}$$

De forma análoga verifica-se que o termo II é dado por:

$$\begin{aligned}
 II = (1 - q) & \left\{ V \left[\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} X_{t-i} | X_{t-d} > r \right] + \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} (1 - \alpha_i^{(2)}) u_{i,d}^{(2)} \right\} + \\
 & + q(1 - q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)^2} u_{i,d}^{(2)^2}.
 \end{aligned} \tag{2.8}$$

Em relação ao termo III , este é dado por

$$\begin{aligned}
 III = 2Cov & \left[\left(I_{t-d}^{(1)} \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} \right), \left(I_{t-d}^{(2)} \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} \right) \right] \\
 = 2 & \left\{ E \left[\left(I_{t-d}^{(1)} \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} \right) \left(I_{t-d}^{(2)} \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} \right) \right] - \right. \\
 & \left. - E \left[I_{t-d}^{(1)} \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} \right] E \left[I_{t-d}^{(2)} \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} \right] \right\} \\
 = -2q(1 - q) & \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i,d}^{(1)} \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i,d}^{(2)}.
 \end{aligned} \tag{2.9}$$

O termo IV toma a forma

$$IV = V \left[\sum_{k=1}^2 Z_t^{(k)} I_{t-d}^{(k)} \right] = q\sigma_Z^{(1)} + (1 - q)\sigma_Z^{(2)}. \tag{2.10}$$

Ao substituir os termos I , II , III e IV apresentados respetivamente pelas equações (2.7), (2.8), (2.9) e (2.10) na equação (2.6) obtem-se facilmente a expressão da variância. Finalmente, para obter a função de autocovariância, tem-se que

$$\begin{aligned}
 \gamma(s) = Cov & \left[X_t, \left(\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t+s-i} + Z_t^{(1)} \right) I_{t+s-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t+s-i} + Z_t^{(2)} \right) I_{t+s-d}^{(2)} \right] \\
 = E & \left[X_t \left(\left(\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t+s-i} + Z_t^{(1)} \right) I_{t+s-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t+s-i} + Z_t^{(2)} \right) I_{t+s-d}^{(2)} \right) \right] \\
 & - E[X_t] E \left[\left(\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t+s-i} + Z_t^{(1)} \right) I_{t+s-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t+s-i} + Z_t^{(2)} \right) I_{t+s-d}^{(2)} \right]
 \end{aligned}$$

$$\begin{aligned}
&= P(X_{t+s-d} \leq r) E \left[X_t \left(\sum_{i=1}^{p(1)} \alpha_i^{(1)} \circ X_{t+s-i} \right) | X_{t+s-d} \leq r \right] + E[X_t Z_t^{(1)} I_{t+s-d}^{(1)}] \\
&\quad + P(X_{t+s-d} > r) E \left[X_t \left(\sum_{i=1}^{p(2)} \alpha_i^{(2)} \circ X_{t+s-i} \right) | X_{t+s-d} > r \right] + E[X_t Z_t^{(2)} I_{t+s-d}^{(2)}] \\
&\quad - E[X_t] E \left[\left(\sum_{i=1}^{p(1)} \alpha_i^{(1)} \circ X_{t+s-i} \right) I_{t+s-d}^{(1)} \right] - E[X_t] E[Z_t^{(1)} I_{t+s-d}^{(1)}] \\
&\quad - E[X_t] E \left[\left(\sum_{i=1}^{p(2)} \alpha_i^{(2)} \circ X_{t+s-i} \right) I_{t+s-d}^{(2)} \right] - E[X_t] E[Z_t^{(2)} I_{t+s-d}^{(2)}].
\end{aligned}$$

Uma vez que X_t é independente de $Z_t^{(k)}$, para $k = 1, 2$ e $s > 0$ tem-se

$$E[X_t Z_t^{(k)} I_{t+s-d}^{(k)}] - E[X_t] E[Z_t^{(k)} I_{t+s-d}^{(k)}] = Cov[X_t, Z_t^{(k)}] = 0, \quad k = 1, 2. \quad (2.11)$$

Então,

$$\begin{aligned}
Cov[X_t, X_{t+s}] &= P(X_{t+s-d} \leq r) E \left[X_t \left(\sum_{i=1}^{p(1)} \alpha_i^{(1)} \circ X_{t+s-i} \right) | X_{t+s-d} \leq r \right] + \\
&\quad + P(X_{t+s-d} > r) E \left[X_t \left(\sum_{i=1}^{p(2)} \alpha_i^{(2)} \circ X_{t+s-i} \right) | X_{t+s-d} > r \right] - \\
&\quad - E[X_t] E \left[\left(\sum_{i=1}^{p(1)} \alpha_i^{(1)} \circ X_{t+s-i} \right) I_{t+s-d}^{(1)} + \left(\sum_{i=1}^{p(2)} \alpha_i^{(2)} \circ X_{t+s-i} \right) I_{t+s-d}^{(2)} \right] \\
&= P(X_{t+s-d} \leq r) E \left[X_t \left(\sum_{i=1}^{p(1)} \alpha_i^{(1)} \circ X_{t+s-i} \right) | X_{t+s-d} \leq r \right] + P(X_{t+s-d} > r) \times \\
&\quad \times E \left[X_t \left(\sum_{i=1}^{p(2)} \alpha_i^{(2)} \circ X_{t+s-i} \right) | X_{t+s-d} > r \right] - E[X_t] \left\{ P(X_{t+s-d} \leq r) \times \right. \\
&\quad \times E \left[\sum_{i=1}^{p(1)} \alpha_i^{(1)} \circ X_{t+s-i} | X_{t+s-d} \leq r \right] + P(X_{t+s-d} > r) E \left[\sum_{i=1}^{p(2)} \alpha_i^{(2)} \circ X_{t+s-i} | X_{t+s-d} > r \right] \Big\} \\
&= q \sum_{i=1}^{p(1)} \alpha_i^{(1)} E[X_t X_{t+s-i} | X_{t+s-d} \leq r] + (1-q) \sum_{i=1}^{p(2)} \alpha_i^{(2)} E[X_t X_{t+s-i} | X_{t+s-d} > r] - \\
&\quad - q \sum_{i=1}^{p(1)} \alpha_i^{(1)} E[X_t] E[X_{t+s-i} | X_{t+s-d} \leq r] - (1-q) \sum_{i=1}^{p(2)} \alpha_i^{(2)} E[X_t] E[X_{t+s-i} | X_{t+s-d} > r]
\end{aligned}$$

$$\begin{aligned}
&= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \left\{ Cov[X_t, X_{t+s-i} | X_{t+s-d} \leq r] + E[X_t | X_{t+s-d} \leq r] E[X_{t+s-i} | X_{t+s-d} \leq r] \right\} \\
&\quad + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \left\{ Cov[X_t, X_{t+s-i} | X_{t+s-d} > r] + E[X_t | X_{t+s-d} > r] E[X_{t+s-i} | X_{t+s-d} > r] \right\} \\
&\quad - q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} E[X_t] E[X_{t+s-i} | X_{t+s-d} \leq r] + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} E[X_t] E[X_{t+s-i} | X_{t+s-d} > r] \\
&= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} Cov[X_t, X_{t+s-i} | X_{t+s-d} \leq r] - q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} E[X_{t+s-i} | X_{t+s-d} \leq r] \times \\
&\quad \times \left\{ E[X_t] - E[X_t | X_{t+s-d} \leq r] \right\} + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} Cov[X_t, X_{t+s-i} | X_{t+s-d} > r] - \\
&\quad - (1-q) E[X_{t+s-i} | X_{t+s-d} > r] \{ E[X_t] - E[X_t | X_{t+s-d} > r] \}.
\end{aligned}$$

Uma vez que

$$E[X_t] = qE[X_t | X_{t-d} \leq r] + (1-q)E[X_t | X_{t-d} > r],$$

então

$$\begin{aligned}
&E[X_t] - E[X_t | X_{t+s-d} \leq r] = \\
&= qE[X_t | X_{t+s-d} \leq r] + (1-q)E[X_t | X_{t+s-d} > r] - E[X_t | X_{t+s-d} \leq r] \\
&= (1-q)(E[X_t | X_{t+s-d} > r] - E[X_t | X_{t+s-d} \leq r]) \\
&= (1-q)(u_{i-s,d}^{(2)} - u_{i-s,d}^{(1)})
\end{aligned}$$

e

$$\begin{aligned}
&E[X_t] - E[X_t | X_{t+s-d} > r] = \\
&= qE[X_t | X_{t+s-d} \leq r] + (1-q)E[X_t | X_{t+s-d} > r] - E[X_t | X_{t+s-d} > r] \\
&= qE[X_t | X_{t+s-d} \leq r] + E[X_t | X_{t+s-d} > r] - qE[X_t | X_{t+s-d} > r] - E[X_t | X_{t+s-d} \leq r] \\
&= q(E[X_t | X_{t+s-d} \leq r] - E[X_t | X_{t+s-d} > r]) \\
&= q(u_{i-s,d}^{(1)} - u_{i-s,d}^{(2)}).
\end{aligned}$$

Assim, tem-se que

$$\begin{aligned}
 Cov[X_t, X_{t+s}] &= \\
 &= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} Cov[X_t, X_{t+s-i} | X_{t+s-d} \leq r] + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} Cov[X_t, X_{t+s-i} | X_{t+s-d} > r] \\
 &\quad + q(1-q) \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i-s, d-s}^{(1)} (u_{0, d-s}^{(1)} - u_{0, d-s}^{(2)}) - q(1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i-s, d-s}^{(2)} (u_{0, d-s}^{(1)} - u_{0, d-s}^{(2)}) \\
 &= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} Cov[X_t, X_{t+s-i} | X_{t+s-d} \leq r] + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} Cov[X_t, X_{t+s-i} | X_{t+s-d} > r] \\
 &\quad + q(1-q) (u_{0, d-s}^{(1)} - u_{0, d-s}^{(2)}) \left\{ \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i-s, d-s}^{(1)} - \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i-s, d-s}^{(2)} \right\} \\
 &= q \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \gamma_{s-i, s-d}^{(1)} + (1-q) \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \gamma_{s-i, s-d}^{(2)} + \\
 &\quad + q(1-q) (u_{0, d-s}^{(1)} - u_{0, d-s}^{(2)}) \left\{ \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} u_{i-s, d-s}^{(1)} - \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} u_{i-s, d-s}^{(2)} \right\}.
 \end{aligned}$$

■

2.3 Métodos de Estimação

2.3.1 Estimadores de Minimos Quadrados Condicionais

Klimko & Nelson (1978) propuseram um método de mínimos quadrados condicionais (MQC) para processos estocásticos. O método consiste em minimizar a soma dos quadrados da diferença de cada observação, X_t , ao seu valor esperado dadas todas as observações anteriores para a obtenção dos estimadores desejados.

Al-Osh & Alzaid (1987), Du & Li (1991), Latour (1998) e Silva (2005b), entre outros autores, fizeram uso deste método para estimar os parâmetros em modelos INAR. Para o modelo SETINAR(2; 1, 1), Monteiro (2010) fez uso de estimadores de mínimos quadrados condicionais para obter as estimativas dos coeficientes $\alpha^{(1)}$, $\alpha^{(2)}$, λ e do limiar r . A seguir, o método MQC é apresentado para o modelo SETINAR(2; $p^{(1)}$, $p^{(2)}$).

O estimador de MQC é o vetor que minimiza a expressão:

$$Q(\boldsymbol{\theta}) := \sum_{t=m+1}^n \left(X_t - g(\boldsymbol{\theta}, X_{t-1}, \dots, X_{t-m}) \right)^2 \equiv \sum_{t=m+1}^n U_t^2, \quad (2.12)$$

onde

$$\boldsymbol{\theta} := \left(\alpha_1^{(1)}, \dots, \alpha_{p^{(1)}}^{(1)}, \lambda^{(1)}, \alpha_1^{(2)}, \dots, \alpha_{p^{(2)}}^{(2)}, \lambda^{(2)} \right), \quad m := \max\{p^{(1)}, p^{(2)}\}$$

e

$$g(\boldsymbol{\theta}, X_{t-1}, \dots, X_{t-m}) := \sum_{i=1}^{p^{(1)}} \left(\alpha_i^{(1)} X_{t-i} I_{t-d}^{(1)} \right) + \sum_{i=1}^{p^{(2)}} \left(\alpha_i^{(2)} X_{t-i} I_{t-d}^{(2)} \right) + \lambda^{(1)} I_{t-d}^{(1)} + \lambda^{(2)} I_{t-d}^{(2)}, \quad (2.13)$$

cuja solução é obtida através da resolução do seguinte sistema

$$\begin{bmatrix} \mathbf{N}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{N}^{(2)} \end{bmatrix} \begin{bmatrix} \alpha_1^{(1)} \\ \vdots \\ \alpha_{p^{(1)}}^{(1)} \\ \lambda^{(1)} \\ \alpha_1^{(2)} \\ \vdots \\ \alpha_{p^{(2)}}^{(2)} \\ \lambda^{(2)} \end{bmatrix} = \begin{bmatrix} S_1^{(1)} \\ \vdots \\ S_{p^{(1)}}^{(1)} \\ S_{\lambda}^{(1)} \\ S_1^{(2)} \\ \vdots \\ S_{p^{(2)}}^{(2)} \\ S_{\lambda}^{(2)} \end{bmatrix}, \quad (2.14)$$

onde $\mathbf{0}$ é uma matriz de dimensão $(p^{(1)} + 1) \times (p^{(2)} + 1)$ de zeros,

$$S_i^{(k)} = \sum_{t=m+1}^n X_t X_{t-i} I_{t-d}^{(k)}, \quad S_{\lambda}^{(k)} = \sum_{t=m+1}^n X_t I_{t-d}^{(k)},$$

e $\mathbf{N}^{(1)}$ e $\mathbf{N}^{(2)}$ são matrizes de dimensão $(p^{(1)} + 1) \times (p^{(1)} + 1)$ e $(p^{(2)} + 1) \times (p^{(2)} + 1)$,

respetivamente, definidas da seguinte maneira

$$\mathbf{N}^{(k)} = \begin{bmatrix} \sum_{t=m+1}^n X_{t-1}^2 I_{t-d}^{(k)} & \sum_{t=m+1}^n X_{t-1} X_{t-2} I_{t-d}^{(k)} & \cdots & \sum_{t=m+1}^n X_{t-1} X_{t-m} I_{t-d}^{(k)} & \sum_{t=m+1}^n X_{t-1} I_{t-d}^{(k)} \\ \sum_{t=m+1}^n X_{t-1} X_{t-2} I_{t-d}^{(k)} & \sum_{t=m+1}^n X_{t-2}^2 I_{t-d}^{(k)} & \ddots & \vdots & \sum_{t=m+1}^n X_{t-2} I_{t-d}^{(k)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{t=m+1}^n X_{t-1} X_{t-m} I_{t-d}^{(k)} & \cdots & \cdots & \sum_{t=m+1}^n X_{t-m}^2 I_{t-d}^{(k)} & \sum_{t=m+1}^n X_{t-m} I_{t-d}^{(k)} \\ \sum_{t=m+1}^n X_{t-1} I_{t-d}^{(k)} & \cdots & \cdots & \sum_{t=m+1}^n X_{t-m}^2 & \sum_{t=m+1}^n I_{t-d}^{(k)} \end{bmatrix},$$

para $k = 1, 2$.

A distribuição assintótica dos estimadores MQC pode ser obtida através dos resultados em Klimko & Nelson (1978).

Proposição 2.1. *Os estimadores MQC são fortemente consistentes e a sua distribuição é assintoticamente normal, i.e.,*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, \mathbf{V}^{-1} \mathbf{W} \mathbf{V}^{-1}), \quad (2.15)$$

onde \mathbf{V} e \mathbf{W} são matrizes quadradas de ordem $p^{(1)} + p^{(2)} + 2$, e definidas da seguinte forma:

$$\mathbf{V} = \begin{bmatrix} q\mathbf{M}^{(1)} & \mathbf{0} \\ \vdots & \vdots \\ \mathbf{0} & (1-q)\mathbf{M}^{(2)} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} q(\mathbf{D}^{(1)} + \mu_Z^{(1)} \mathbf{M}^{(1)}) & \mathbf{0} \\ \vdots & \vdots \\ \mathbf{0} & (1-q)(\mathbf{D}^{(2)} + \mu_Z^{(2)} \mathbf{M}^{(2)}) \end{bmatrix}, \quad (2.16)$$

com

$$\mathbf{M}^{(k)} = \begin{bmatrix} m_1^{(k)} & m_{12}^{(k)} & \dots & m_{1p^{(k)}}^{(k)} & u_{1,d}^{(k)} \\ m_{12}^{(k)} & m_2^{(k)} & \dots & m_{2p^{(k)}}^{(k)} & u_{2,d}^{(k)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{1p^{(k)}}^{(k)} & m_{2p^{(k)}}^{(k)} & \dots & m_{p^{(k)}}^{(k)} & u_{p^{(k)},d}^{(k)} \\ u_{1,d}^{(k)} & u_{2,d}^{(k)} & \dots & u_{p^{(k)},d}^{(k)} & 1 \end{bmatrix}, \quad k = 1, 2,$$

onde $m_i^{(1)} := E[X_{t-i}^2 | X_{t-d} \leq r]$, $m_i^{(2)} = m_{ii}^{(2)} := E[X_{t-i}^2 | X_{t-d} > r]$, $m_{ij}^{(1)} := E[X_{t-i}X_{t-j} | X_{t-d} \leq r]$ e $m_{ij}^{(2)} := E[X_{t-i}X_{t-j} | X_{t-d} > r]$. Finalmente,

$$\mathbf{D}^{(k)} = \begin{bmatrix} \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) w_{l11}^{(k)} & \dots & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) w_{l1p^{(k)}}^{(k)} & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) m_{l1}^{(k)} \\ \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) w_{l12}^{(k)} & \dots & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) w_{l2p^{(k)}}^{(k)} & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) m_{l2}^{(k)} \\ \vdots & \dots & \vdots & \vdots \\ \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) w_{lp^{(k)}p^{(k)}}^{(k)} & \dots & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) w_{lp^{(k)}p^{(k)}}^{(k)} & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) m_{lp^{(k)}}^{(k)} \\ \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) m_{l1}^{(k)} & \dots & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) m_{lp^{(k)}}^{(k)} & \sum_{l=1}^{p^{(k)}} \alpha_l^{(k)} (1 - \alpha_l^{(k)}) u_{l,d}^{(k)} \end{bmatrix}, \quad k = 1, 2,$$

com $w_{ijl}^{(1)} := E[X_{t-i}X_{t-j}X_{t-l} | X_{t-d} \leq r]$ e $w_{ijl}^{(2)} := E[X_{t-i}X_{t-j}X_{t-l} | X_{t-d} > r]$, para $l, i, j = 1, \dots, p^{(k)}$.

Demonstração. É simples verificar que a função g e as derivadas parciais

$$\frac{\partial g(\boldsymbol{\theta}, X_{t-1}, \dots, X_{t-m})}{\partial \theta_i}, \frac{\partial^2 g(\boldsymbol{\theta}, X_{t-1}, \dots, X_{t-m})}{\partial \theta_i \partial \theta_j}, \frac{\partial^3 g(\boldsymbol{\theta}, X_{t-1}, \dots, X_{t-m})}{\partial \theta_i \partial \theta_j \partial \theta_k},$$

com $i, j, k = 1, \dots, (p^{(1)} + p^{(2)} + 2)$ satisfazem todas as condições de regularidade do Teorema 3.1 de Klimko & Nelson (1978). Desta forma os estimadores de mínimos quadrados condicionais para um processo SETINAR(2; 1, 1) são fortemente consistentes. A prova da normalidade assintótica dos estimadores obtém-se a partir dos argumentos utilizados na demonstração do Teorema 3.1 em Monteiro *et al.* (2012). ■

2.3.2 Verosimilhança Condicional

Para se obter a verosimilhança condicional L do modelo **SETINAR** é suficiente conhecer as probabilidades de transição uma vez que

$$L\left(\alpha_1^{(1)}, \dots, \alpha_{p^{(1)}}^{(1)}, \alpha_1^{(2)}, \dots, \alpha_{p^{(2)}}^{(2)}, \lambda^{(1)}, \lambda^{(2)}\right) = \prod_{t=m+1}^n P(X_t | X_{t-1}, \dots, X_{t-m}),$$

onde $m := \max(p^{(1)}, p^{(2)})$. Adaptando ao modelo **SETINAR**(2; $p^{(1)}, p^{(2)}$) o procedimento de Bu *et al.* (2008) usado no modelo **INAR**(p), verifica-se que

$$X_t = \left(\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} + Z_t^{(1)} \right) I_{t-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} + Z_t^{(2)} \right) I_{t-d}^{(2)},$$

pode ser visto como a convolução de

$$\left(\alpha_1^{(1)} \circ X_{t-1} \right) I_{t-d}^{(1)} + \left(\alpha_1^{(2)} \circ X_{t-1} \right) I_{t-d}^{(2)},$$

com

$$\left(\sum_{i=2}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} + Z_t^{(1)} \right) I_{t-d}^{(1)} + \left(\sum_{i=2}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} + Z_t^{(2)} \right) I_{t-d}^{(2)}.$$

Analogamente,

$$\left(\sum_{i=2}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} + Z_t^{(1)} \right) I_{t-d}^{(1)} + \left(\sum_{i=2}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} + Z_t^{(2)} \right) I_{t-d}^{(2)},$$

pode ser interpretado como a convolução entre

$$\left(\alpha_2^{(1)} \circ X_{t-2} \right) I_{t-d}^{(1)} + \left(\alpha_2^{(2)} \circ X_{t-2} \right) I_{t-d}^{(2)},$$

com

$$\left(\sum_{i=3}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} + Z_t^{(1)} \right) I_{t-d}^{(1)} + \left(\sum_{i=3}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} + Z_t^{(2)} \right) I_{t-d}^{(2)}$$

e assim sucessivamente.

No caso em que as inovações têm distribuição de Poisson, vem a seguinte expressão para as probabilidades de transição

$$\begin{aligned}
 P(X_t | X_{t-1}, \dots, X_{t-m}) &= P(X_t = x_t | X_{t-1} = x_{t-1}, \dots, X_{t-m} = x_{t-m}) \\
 &= P \left(\left[\sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i} + Z_t^{(1)} \right] I_{t-d}^{(1)} + \left[\sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i} + Z_t^{(2)} \right] I_{t-d}^{(2)} = x_t \middle| x_{t-1}, \dots, x_{t-m} \right) \\
 &= \sum_{k=1}^2 \left\{ \left[\sum_{i_1=0}^{\min(x_t, x_{t-1})} C_{i_1}^{x_{t-1}} (\alpha_1^{(k)})^{i_1} (1 - \alpha_1^{(k)})^{x_{t-1}-i_1} \times \right. \right. \\
 &\quad \times \left[\sum_{i_2=0}^{\min(x_{t-2}, x_{t-1}-i_1)} C_{i_2}^{x_{t-2}} (\alpha_2^{(k)})^{i_2} (1 - \alpha_2^{(k)})^{x_{t-2}-i_2} \dots \times \right. \\
 &\quad \times \left[\sum_{i_m=0}^{\min(x_{t-m}, x_{t-1}-i_1-\dots-i_{m-1})} C_{i_m}^{x_{t-m}} (\alpha_m^{(k)})^{i_m} (1 - \alpha_m^{(k)})^{x_{t-m}-i_m} \times \right. \\
 &\quad \left. \left. \left. \times e^{-\lambda^{(k)}} \frac{(\lambda^{(k)})^{x_{t-1}-i_1-\dots-i_{m-1}}}{(x_{t-1}-i_1-\dots-i_{m-1})!} \right] \dots \right] \right] \left. \right\}.
 \end{aligned}$$

Com vista a facilitar o processo de estimação de parâmetros no modelo **SETINAR**, faz-se uma ampliação de dados, tal como foi efetuado para um modelo **INARMA** por Neal & Subba Rao (2007) onde os dados ampliados são tratados como variáveis latentes a serem atualizadas. Estimativas de máxima verosimilhança (EMV) em dados ampliados são realizadas, em geral, através do uso do algoritmo de ampliação de dados EM proposto por Dempster *et al.* (1977).

O procedimento para imputação de variáveis latentes no modelo **SETINAR**(2; $p^{(1)}, p^{(2)}$), o qual segue o procedimento eficiente utilizado por Neal & Subba Rao (2007) para a obtenção da verosimilhança e para a atualização dos dados ampliados, é descrito a seguir.

Seja a série temporal observada $\mathbf{x} := (x_{1-m}, x_{2-m}, \dots, x_n)$ onde $m := \max(p^{(1)}, p^{(2)})$ representa a ordem máxima dos processos **INAR**(p) intervenientes. Seja $Y_{t,i}^{(k)} := \alpha_i^{(k)} \circ X_{t-i}$ e $Y_t^{(k)} = \sum_{i=1}^{p^{(k)}} Y_{t,i}^{(k)}$, para $t \in \mathbb{N}$, $i = 1, \dots, p^{(k)}$ e $k = 1, 2$. Então

$Z_t^{(k)} = X_t - Y_t^{(k)}$, $t \in \mathbb{N}$, $k = 1, 2$. Isto significa que

$$\begin{cases} Y_t^{(1)} = \sum_{i=1}^{p^{(1)}} \alpha_i^{(1)} \circ X_{t-i}, & X_{t-d} \leq r \\ Y_t^{(2)} = \sum_{i=1}^{p^{(2)}} \alpha_i^{(2)} \circ X_{t-i}, & X_{t-d} > r \end{cases}$$

e

$$\begin{cases} Z_t^{(1)} = X_t - \sum_{i=1}^{p^{(1)}} Y_{t,i}^{(1)}, & X_{t-d} \leq r \\ Z_t^{(2)} = X_t - \sum_{i=1}^{p^{(2)}} Y_{t,i}^{(2)}, & X_{t-d} > r \end{cases}.$$

Definam-se ainda $\mathbf{Y} := (Y_1, \dots, Y_n)$, sendo $Y_t = Y_t^{(1)} I_{t-d}^{(1)} + Y_t^{(2)} I_{t-d}^{(2)}$ e $\mathbf{Z} := (Z_1, \dots, Z_n)$, onde $Z_t = Z_t^{(1)} I_{t-d}^{(1)} + Z_t^{(2)} I_{t-d}^{(2)}$. Em suma, considera-se como informação extra, além da fornecida pelos dados \mathbf{x} , a fornecida por $(y_1^{(k)}, z_1^{(k)}), \dots, (y_n^{(k)}, z_n^{(k)})$. Uma vez que os $Y_{t,1}^{(k)}, Y_{t,2}^{(k)}, \dots, Y_{t,p}^{(k)}$ são independentes e independentes de $Z_t^{(k)}$, a verossimilhança das variáveis latentes dados os valores observados \mathbf{x} e os parâmetros $\boldsymbol{\theta} := (\boldsymbol{\alpha}, \boldsymbol{\lambda})$ pode ser escrita como

$$f(y_t, z_t | x_{t-1}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = f(y_t | x_{t-1}, \boldsymbol{\alpha}) f(z_t | \boldsymbol{\lambda}), \quad (2.17)$$

onde $\boldsymbol{\alpha} := (\alpha^{(1)}, \alpha^{(2)})$ e $\boldsymbol{\lambda} := (\lambda^{(1)}, \lambda^{(2)})$, sendo

$$\begin{cases} \alpha^{(1)} := (\alpha_1^{(1)}, \alpha_2^{(1)}, \dots, \alpha_{p^{(1)}}^{(1)}) & e \quad \lambda^{(1)} \quad se \quad x_{t-d} \leq r \\ \alpha^{(2)} := (\alpha_1^{(2)}, \alpha_2^{(2)}, \dots, \alpha_{p^{(2)}}^{(2)}) & e \quad \lambda^{(2)} \quad se \quad x_{t-d} > r \end{cases}.$$

A representação (2.17), que define a distribuição condicional, calculada no instante $1 \leq t \leq n$, é dada por

$$f(y_t, z_t | \mathbf{x}_{t-1}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = \sum_{k=1}^2 f(y_t^{(k)} | \mathbf{x}_{t-1}, \boldsymbol{\alpha}^{(k)}) f(z_t^{(k)} | \lambda_t^{(k)}) I_{t-d}^{(k)}$$

$$\begin{aligned}
&= \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k) z_t^{(k)}}}{z_t^{(k)}!} f(y_t^{(k)} | \mathbf{x}_{t-1}, \alpha^{(k)}) \\
&= \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k) z_t^{(k)}}}{z_t^{(k)}!} f(y_{t,1}^{(k)}, y_{t,2}^{(k)}, \dots, y_{t,p^{(k)}}^{(k)} | \mathbf{x}_{t-1}, \alpha^{(k)}) \\
&= \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k) z_t^{(k)}}}{z_t^{(k)}!} f(y_{t,1}^{(k)} | \mathbf{x}_{t-1}, \alpha_1^{(k)}) \times \dots \times f(y_{t,p^{(k)}}^{(k)} | \mathbf{x}_{t-p^{(k)}}, \alpha_{p^{(k)}}^{(k)}) \\
&= \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k) z_t^{(k)}}}{z_t^{(k)}!} \binom{x_{t-1}}{y_{t,1}^{(1)}} \alpha_1^{(k) y_{t,1}^{(k)}} (1 - \alpha_1^{(k)})^{x_{t-1} - y_{t,1}^{(k)}} \times \\
&\quad \times \dots \times \binom{x_{t-p^{(k)}}}{y_{t,p^{(k)}}^{(1)}} \alpha_p^{(k) y_{t,p^{(k)}}^{(k)}} (1 - \alpha_p^{(k)})^{x_{t-p^{(k)}} - y_{t,p^{(k)}}^{(k)}} \\
&= \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k) z_t^{(k)}}}{z_t^{(k)}!} \prod_{i=1}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} \alpha_i^{(k) y_{t,i}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i} - y_{t,i}^{(k)}},
\end{aligned}$$

com $x_t = \left(\sum_{i=1}^{p^{(1)}} y_{t,i}^{(1)} + z_t^{(1)} \right) I_{t-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}} y_{t,i}^{(2)} + z_t^{(2)} \right) I_{t-d}^{(2)}$ e $y_{t,i}^{(k)} \leq x_{t-i}$, para $i = 1, \dots, p^{(k)}, k = 1, 2$. Assim, a distribuição conjunta de (\mathbf{Y}, \mathbf{Z}) condicional às observações $\mathbf{x} := \{x_{1-m}, x_{2-m}, \dots, x_1, \dots, x_n\}$ e $\boldsymbol{\theta} := (\boldsymbol{\alpha}, \boldsymbol{\lambda})$ é

$$L(\mathbf{y}, \mathbf{z}; \boldsymbol{\theta}) = f(\mathbf{y}, \mathbf{z} | \mathbf{x}, \boldsymbol{\theta}) = \prod_{t=1}^n \sum_{k=1}^2 \left\{ I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k) z_t^{(k)}}}{z_t^{(k)}!} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} (\alpha_i^{(k)})^{y_{t,i}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i} - y_{t,i}^{(k)}} \right\}$$

restrito a

$$x_t = \left(\sum_{i=1}^{p^{(1)}} y_{t,i}^{(1)} + z_t^{(1)} \right) I_{t-d}^{(1)} + \left(\sum_{i=1}^{p^{(2)}} y_{t,i}^{(2)} + z_t^{(2)} \right) I_{t-d}^{(2)}.$$

Algoritmo EM de Monte Carlo

Dado um vetor de estimativas iniciais $\boldsymbol{\theta}^{(0)}$, os dados observados \mathbf{x} e os dados ampliados $\mathbf{w} = (\mathbf{y}, \mathbf{z})$, para se obter as estimativas de ML através do algoritmo MCEM consideram-se dois passos como apresentado na Secção 1.5.3: o passo E - *Expectation* e o passo M - *Maximization*.

No passo E o algoritmo calcula uma aproximação do logaritmo da função de verosimilhança, $l(\theta) := \ln [L(\mathbf{w}; \theta) | \mathbf{x}, \theta^{(t)}]$ onde

$$\begin{aligned} l(\theta) &= \sum_{t=1}^n \ln \left\{ \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k)z_t^{(k)}}}{z_t^{(k)}!} \prod_{i=1}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} \alpha_i^{(k)y_{t,i}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i}-y_{t,i}^{(k)}} \right\} \\ &= \sum_{t=1}^n \ln \left\{ \underbrace{I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}}{z_t^{(1)}!} \prod_{i=1}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}_{l_1} + \right. \\ &\quad \left. + \underbrace{I_{t-d}^{(2)} \frac{e^{(-\lambda^{(2)})} \lambda^{(2)z_t^{(2)}}}{z_t^{(2)}!} \prod_{i=1}^{p^{(2)}} \binom{x_{t-i}}{y_{t,i}^{(2)}} \alpha_i^{(2)y_{t,i}^{(2)}} (1 - \alpha_i^{(2)})^{x_{t-i}-y_{t,i}^{(2)}}}_{l_2} \right\}. \end{aligned}$$

Essa aproximação é dada por $Q(\theta | \theta^{(t)}) = \frac{1}{m} \sum_{j=1}^m \ln [L(\mathbf{w}_j; \theta) | \mathbf{x}, \theta^{(t)}]$. Seguidamente o passo M consiste em maximizar $Q(\theta | \theta^{(t)})$, isto é, em determinar $\hat{\theta} = \mathbf{argmax} \quad Q(\theta | \theta^{(t)})$. A derivação parcial de Q em relação aos parâmetros $\boldsymbol{\alpha}^{(k)}$ e $\lambda^{(k)}$, com $k = 1, 2$ e $i = \{1, \dots, p^{(k)}\}$, resulta em

$$\left\{ \begin{aligned} \frac{\partial Q}{\partial \alpha_i^{(k)}} &= \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \frac{y_{t,i}^{(k)} - \alpha_i^{(k)} x_{t-i}}{\alpha_i^{(k)} (1 - \alpha_i^{(k)})} I_{t-d}^{(k)} \\ \frac{\partial Q}{\partial \lambda^{(k)}} &= \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \left\{ \left(-1 + \frac{z_t^{(k)}}{\lambda^{(k)}} \right) \times I_{t-d}^{(k)} \right\} \end{aligned} \right\},$$

conduzindo a

$$\left\{ \begin{array}{l} \hat{\alpha}_i^{(k)} = \frac{\sum_{j=1}^m \sum_{t=1}^n y_{t,i}^{(k)} I_{t-d}^{(k)}}{\sum_{j=1}^m \sum_{t=1}^n x_{t-i} I_{t-d}^{(k)}} \\ \hat{\lambda}^{(k)} = \frac{\sum_{j=1}^m \sum_{t=1}^n z_t^{(k)} I_{t-d}^{(k)}}{\sum_{j=1}^m \sum_{t=1}^n I_{t-d}^{(k)}} \end{array} \right. .$$

Esta metodologia pode ser implementada computacionalmente através dos seguintes algoritmos:

Algoritmo 2.1: Algoritmo EM -SETINAR(2; $p^{(1)}, p^{(2)}$) - Passo E

Entrada: $\theta^{(0)}$, \mathbf{x} , m

Saída: $Q(\theta|\theta^{(t)}) = \frac{1}{m} \sum_{j=1}^m \ln[L(\mathbf{w}_j; \theta)|\mathbf{x}, \theta^{(t)}]$

inicio

para $j \leftarrow 1$ **até** m **fazer**

 Gerar $Y_j^{(k)}$

 Calcular $z_j^{(k)}$

$Q_j \leftarrow \ln[L(\mathbf{w}_j; \theta)|\mathbf{x}, \theta^{(t)}]$

fin para

$Q \leftarrow \frac{1}{m} \sum_{j=1}^m Q_j$

fin

2.3.3 Inferência bayesiana

O problema da estimação bayesiana dos parâmetros pode ser analisado considerando a situação de as ordens $p^{(1)}$ e $p^{(2)}$ dos processos INAR de cada um dos regimes serem conhecidos ou desconhecidos.

Algoritmo 2.2: Algoritmo EM -SETINAR(2; $p^{(1)}, p^{(2)}$) - Passo M

Entrada: Q
Saída: $\hat{\theta} = \operatorname{argmax} Q$
inicio
 para $k \leftarrow 1$ **até** 2 **hacer**
 para $p \leftarrow 1$ **até** p_k **hacer**
 Calcular $\frac{\partial Q}{\partial \alpha_p^{(k)}} = 0$
 fin para
 Calcular $\frac{\partial Q}{\partial \lambda^{(k)}} = 0$
 fin para
 $\hat{\theta} = \operatorname{argmax} Q$
fin

No caso de $p^{(1)}$ e $p^{(2)}$ serem conhecidos, o problema resolve-se usando o método MCMC, onde para cada regime k ($k = 1, 2$) os parâmetros $\alpha^{(k)}$ e $\lambda^{(k)}$ são estimados supondo a ordem fixa $p^{(k)}$. O movimento que lhe está associado designa-se movimento dentro do modelo.

No caso das ordens serem desconhecidas aplica-se o algoritmo *Reversible Jump Markov Chain Monte Carlo* - RJMCMC - proposto por Green (1995) para que, além de se estimarem os parâmetros se possa também estimar as ordens dos modelos de cada um dos regimes. O procedimento RJMCMC possibilita que o amostrador efetue saltos entre espaços de dimensões diferentes. Neste caso, a estimativa é feita em duas etapas. A primeira consiste no movimento dentro do modelo, o que corresponde a obter estimativas bayesianas para os parâmetros considerando as ordens fixas. Seguidamente estima-se a nova ordem usando o procedimento RJMCMC, proposto por Enciso-Mora *et al.* (2008), designado por movimento entre modelos

Com o objetivo de se obterem estimativas bayesianas, considera-se o procedimento anteriormente descrito de ampliação de dados. Consequentemente para determinar a verosimilhança conjunta total de $(\mathbf{y}_t, \mathbf{z}_t)$ condicionado aos dados \mathbf{x} é necessário definir as distribuições *a priori* do conjunto de parâmetros $\boldsymbol{\theta} := (\boldsymbol{\alpha}, \boldsymbol{\lambda})$ com $\boldsymbol{\alpha} := (\alpha^{(1)}, \alpha^{(2)})$ e $\boldsymbol{\lambda} := (\lambda^{(1)}, \lambda^{(2)})$. Optou-se por distribuições *a priori* tal como definidas em Neal

& Subba Rao (2007): *a priori* não informativas uniformes e restritas ao espaço do parâmetro α e uma distribuição gama conjugada para λ .

Seja $k = 1, 2$ os ramos do modelo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$, as distribuições *a priori* são definidas como

$$\begin{aligned} \pi \left(\alpha_i^{(k)} | p^{(k)} \right) &\propto 1, \quad 0 \leq \alpha_i^{(k)} < 1, \quad 1 \leq i \leq p^{(k)}, \quad \sum_{i=1}^{p^{(k)}} \alpha_i^{(k)} < 1; \\ \pi \left(\lambda^{(k)} \right) &= \text{Gama}(A, B), \quad \text{para } A, B > 0. \end{aligned}$$

Supondo, numa primeira fase, que as ordens $p^{(1)}$ e $p^{(2)}$ são conhecidas a verosimilhança conjunta condicional a \mathbf{x} é dada por

$$\begin{aligned} f(\mathbf{y}, \mathbf{z}, \alpha, \lambda | \mathbf{x}) &= L(\mathbf{y}_t, \mathbf{z}_t; \theta) \prod_{k=1}^2 \pi \left(\alpha^{(k)} \right) \pi \left(\lambda^{(k)} \right) \\ &\propto \prod_{t=1}^n \sum_{k=1}^2 \left\{ I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} [\lambda^{(k)}]^{z_t^{(k)}}}{z_t^{(k)}!} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} \alpha_i^{(k) y_{t,i}^{(k)}} \left(1 - \alpha_i^{(k)} \right)^{x_{t-i} - y_{t,i}^{(k)}} \right\} \times \\ &\quad \times (\lambda^{(1)})^{A^{(1)}} e^{-B^{(1)} \lambda^{(1)}} \times (\lambda^{(2)})^{A^{(2)}} e^{-B^{(2)} \lambda^{(2)}}. \end{aligned} \tag{2.18}$$

A distribuição condicional completa para o parâmetro $\alpha_i^{(k)}$ com $k = 1, 2$ e $i = 1, \dots, p^{(k)}$, obtem-se a partir da equação (2.18) e é dada por

$$\pi \left(\alpha_i^{(k)} \middle| \mathbf{y}, \mathbf{z}, \lambda, \mathbf{x}, \alpha_{-i}^{(k)} \right) \propto \left(\alpha_i^{(k)} \right)^{\sum_{t=1}^n y_{t,i}^{(k)} I_{t-d}^{(k)}} \left(1 - \alpha_i^{(k)} \right)^{\sum_{t=1}^n (x_{t-i} - y_{t,i}^{(k)}) I_{t-d}^{(k)}},$$

isto é,

$$\alpha_i^{(k)} \middle| \mathbf{y}, \mathbf{z}, \lambda, \mathbf{x}, \alpha_{-i}^{(k)} \sim \text{Beta} \left(1 + \sum_{t=1}^n y_{t,i}^{(k)} I_{t-d}^{(k)}, 1 + \sum_{t=1}^n (x_{t-i} - y_{t,i}^{(k)}) I_{t-d}^{(k)} \right). \tag{2.19}$$

Analogamente, a partir da equação (2.18), atendendo a que

$$\pi \left(\lambda_t^{(k)} \middle| \mathbf{y}, \mathbf{z}, \mathbf{x}, \alpha^{(k)} \right) \propto (\lambda^{(k)})^{A^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} z_t^{(k)}} e^{-\lambda^{(k)} [B^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)}]},$$

isto é,

$$\lambda_t^{(1)} \Big| \mathbf{y}, \mathbf{z}, \mathbf{x}, \alpha^{(k)} \sim Gama \left(1 + A^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} z_t^{(k)}, B^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} \right).$$

Movimento dentro de um modelo

Dispondo dos dados observados \mathbf{x} e dos aumentados (\mathbf{y}, \mathbf{z}) , ter-se-á de proceder um processo de atualização sucessivo dos parâmetros e dos dados aumentados.

A atualização de cada parâmetro $\alpha_i^{(k)}$ é feita a partir da distribuição condicional completa

$$\alpha_i^{(k)} \Big| \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \mathbf{x}, \boldsymbol{\alpha}_{-i}^{(k)} \sim Beta \left(1 + \sum_{t=1}^n y_{t,i}^{(k)} I_{t-d}^{(k)}, 1 + \sum_{t=1}^n \left(x_{t-i} - y_{t,i}^{(k)} \right) I_{t-d}^{(k)} \right),$$

sujeita à restrição $\sum_{i=1}^{p^{(k)}} \alpha_i^{(k)} < 1$.

Os parâmetros $\alpha^{(k)}$ são atualizados de forma independente para os diferentes valores de k . Quando $x_{t-d} \leq r$ é atribuído $k = 1$ e atualiza-se $\alpha^{(1)}$ caso contrário $k = 2$ e os valores de $\alpha^{(2)}$ é que são atualizados. Amostra-se $\alpha^{(k)}$ de acordo com o procedimento abaixo:

- (i) Gerar o novo $\alpha_i^{'(k)} \sim Beta$ dada na equação (2.19);
- (ii) Se $\sum_{i=1}^{p^{(k)}} \alpha_i^{'(k)} < 1$ então aceita $\alpha_i^{'(k)}$ senão repete o passo (i).

A atualização de λ é feita gerando o valor λ da distribuição condicional

$$\lambda^{(k)} \Big| \mathbf{y}, \mathbf{z}, \mathbf{x}, \alpha^{(k)} \sim Gama \left(1 + A^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} z_t^{(k)}, B^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} \right),$$

para $i = 1, \dots, p^{(k)}$ e $k = 1, 2$. Os parâmetros $\lambda^{(1)}$ e $\lambda^{(2)}$ são atualizados de forma independente. Analogamente quando $x_{t-d} \leq r$ é atribuído $k = 1$ e amostra-se o novo $\lambda^{'(1)}$ caso contrário $k = 2$ e o novo $(\lambda^{'})^{(2)}$ é amostrado. Em cada instante t , as

componentes (\mathbf{y}, \mathbf{z}) são atualizadas da seguinte forma:

- (i) Para $1 \leq t \leq n$
 - * Se $x_{t-d} \leq r$ então $k \leftarrow 1$;
 - * Senão $k \leftarrow 2$.
- (ii) Para $1 \leq i \leq p^{(k)}$, gerar $y'_{t,i} \sim \text{Bin} \left(x_{t-i}, \alpha_i^{(k)} \right)$;
- (iii) Se $x_t < \sum_{i=1}^{p^{(k)}} y'_{t,i}$ senão repetir [ii];
- (iv) Definir $z'_t = x_t - \sum_{i=1}^{p^{(k)}} y'_{t,i}$;
- (iv) Aceitar o movimento com probabilidade $A = 1 \wedge \left(\lambda^{(k)} \right)^{Z_t'^{(k)} - Z_t^{(k)}} \frac{Z_t^{(k)}!}{Z_t'^{(k)}!}$, se mudança aceite, toma-se $(\mathbf{y}_t, \mathbf{z}_t) = (\mathbf{y}'_t, \mathbf{z}'_t)$ caso contrário mantém-se $(\mathbf{y}_t, \mathbf{z}_t)$.

Mostremos que a probabilidade de aceitação do movimento é dada por

$$A = 1 \wedge \left(\lambda^{(k)} \right)^{Z_t'^{(k)} - Z_t^{(k)}} \frac{Z_t^{(k)}!}{Z_t'^{(k)}!}. \quad (2.20)$$

De acordo com o algoritmo Metropolis-Hastings, usando o amostrador independente, a probabilidade de aceitação de A de cada par (y'_t, z'_t) é dada por

$$A = 1 \wedge \underbrace{\frac{f(y'_t, z'_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda})}{f(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda})}}_I \times \underbrace{\frac{q(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha})}{q(y'_t, z'_t | \mathbf{x}, \boldsymbol{\alpha})}}_{II}, \quad (2.21)$$

onde $q(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha})$ é a distribuição condicional proponente para (y_t, z_t) . A distribuição proponente q foi escolhida de forma que a maioria dos termos de (2.21) fossem cancelados.

O termo I da probabilidade de aceitação descrita na equação (2.20) é dado por

$$\frac{f(y'_t, z'_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda})}{f(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda})},$$

onde,

$$f(y'_t, z'_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{[-\lambda^{(k)}]} \lambda^{(k)z'_t}}{z'_t{}^{(k)}!} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y'_{t,i}{}^{(k)}} \alpha_i^{(k)y'_{t,i}{}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i}-y'_{t,i}{}^{(k)}}$$

e

$$f(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = \sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{[-\lambda^{(k)}]} \lambda^{(k)z_t}}{z_t{}^{(k)}!} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} \alpha_i^{(k)y_{t,i}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i}-y_{t,i}^{(k)}}.$$

O termo II da probabilidade de aceitação é definido pela expressão

$$\frac{q(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha})}{q(y'_t, z'_t | \mathbf{x}, \boldsymbol{\alpha})},$$

com

$$q(y_t, z_t | \mathbf{x}, \boldsymbol{\alpha}) = \sum_{k=1}^2 I_{t-d}^{(k)} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} \alpha_i^{(k)y_{t,i}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i}-y_{t,i}^{(k)}}$$

e

$$q(y'_t, z'_t | \mathbf{x}, \boldsymbol{\alpha}) = \sum_{k=1}^2 I_{t-d}^{(k)} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y'_{t,i}{}^{(k)}} \alpha_i^{(k)y'_{t,i}{}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i}-y'_{t,i}{}^{(k)}}.$$

Assim quando estamos a estimar os parâmetros do primeiro regime, isto é, quando

$x_{t-d} \leq r$ os termos I e II da probabilidade de aceitação são dados por

$$\frac{f(\mathbf{y}'_t, \mathbf{z}'_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda})}{f(\mathbf{y}_t, \mathbf{z}_t | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\lambda})} = \frac{\frac{e^{[-\lambda^{(1)}]} \lambda^{(1)z'_t}}{z'_t{}^{(1)}!} \prod_{i=0}^{p^{(1)}} \binom{x_{t-i}}{y'_{t,i}{}^{(1)}} \alpha_i^{(1)y'_{t,i}{}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y'_{t,i}{}^{(1)}}}{\frac{e^{[-\lambda^{(1)}]} \lambda^{(1)z_t}}{z_t{}^{(1)}!} \prod_{i=0}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}$$

$$\frac{q(\mathbf{y}_t, \mathbf{z}_t | \mathbf{x}, \alpha)}{q(\mathbf{y}'_t, \mathbf{z}'_t | \mathbf{x}, \alpha)} = \frac{\prod_{i=0}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}{\prod_{i=0}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}'^{(1)}} \alpha_i^{(1)y_{t,i}'^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}'^{(1)}}}.$$

De forma análoga, quando os parâmetros a estimar são os relativos ao segundo regime, quando $x_{t-d} > r$, os termos *I* e *II* da probabilidade de aceitação são determinados por

$$\frac{f(\mathbf{y}'_t, \mathbf{z}'_t | \mathbf{x}, \alpha, \lambda)}{f(\mathbf{y}_t, \mathbf{z}_t | \mathbf{x}, \alpha, \lambda)} = \frac{\frac{e^{-\lambda^{(2)}} \lambda^{(2)z_t'} p^{(2)}}{z_t'^{(2)}!} \prod_{i=0}^{p^{(2)}} \binom{x_{t-i}}{y_{t,i}'^{(2)}} \alpha_i^{(2)y_{t,i}'^{(2)}} (1 - \alpha_i^{(2)})^{x_{t-i}-y_{t,i}'^{(2)}}}{\frac{e^{-\lambda^{(2)}} \lambda^{(2)z_t} p^{(2)}}{z_t^{(2)}!} \prod_{i=0}^{p^{(2)}} \binom{x_{t-i}}{y_{t,i}^{(2)}} \alpha_i^{(2)y_{t,i}^{(2)}} (1 - \alpha_i^{(2)})^{x_{t-i}-y_{t,i}^{(2)}}}$$

$$\frac{q(\mathbf{y}_t, \mathbf{z}_t | \mathbf{x}, \alpha)}{q(\mathbf{y}'_t, \mathbf{z}'_t | \mathbf{x}, \alpha)} = \frac{\prod_{i=0}^{p^{(2)}} \binom{x_{t-i}}{y_{t,i}^{(2)}} \alpha_i^{(2)y_{t,i}^{(2)}} (1 - \alpha_i^{(2)})^{x_{t-i}-y_{t,i}^{(2)}}}{\prod_{i=0}^{p^{(2)}} \binom{x_{t-i}}{y_{t,i}'^{(2)}} \alpha_i^{(2)y_{t,i}'^{(2)}} (1 - \alpha_i^{(2)})^{x_{t-i}-y_{t,i}'^{(2)}}}.$$

Assim a probabilidade de aceitação se resume em

$$\begin{cases} A = 1 \wedge \lambda^{(1)[z_t^{(1)} - z_t'^{(1)}]} \frac{z_t^{(1)}!}{z_t'^{(1)}!} & \text{se } x_{t-d} \leq r, \\ A = 1 \wedge \lambda^{(2)[z_t'^{(2)} - z_t^{(2)}]} \frac{z_t'^{(2)}!}{z_t^{(2)}!} & \text{se } x_{t-d} > r. \end{cases}$$

Movimento entre modelos

No caso em que $p^{(1)}$ e $p^{(2)}$ são desconhecidos, para estimar estas ordens e seguindo a metodologia bayesiana ter-se-á de considerar distribuições *a priori* para estes parâmetros e de usar o procedimento que permitam mudanças entre modelos.

Consideram-se distribuições *a priori* não informativas para $p^{(k)}$ e distribuições

uniformes também não informativas para $\alpha|p^{(k)}$ dadas por

$$\pi(\alpha|p^{(k)}) \propto p^{(k)}!, \quad 0 < \alpha_i^{(k)}, \quad i = 1, \dots, p^{(k)}, \quad \sum_{i=1}^{p^{(k)}} \alpha_i^{(k)} < 1;$$

$$\pi(\lambda) \propto e^{-A\lambda^{(k)}} (\lambda^{(k)})^{B-1}, \quad \lambda^{(k)} > 0 \text{ e } A, B > 0.$$

Para obter $f((y), (z), (\alpha), (\lambda), p^{(1)}, p^{(2)} | \mathbf{x})$ basta multiplicar a expressão (2.18) pelo fator $p^{(1)}!p^{(2)}!\pi(p^{(1)})\pi(p^{(2)})$ sujeitas a $\sum_{i=1}^m y_{t,i} + z_t = x_t, 1 \leq t \leq n$ e $m := \max(p^{(1)}, p^{(2)})$.

As atualizações de $\alpha_i^{(k)}$ e $\lambda^{(k)}$ são feitas da forma anteriormente indicada, uma vez que as distribuições condicionais completas não se alteram.

Para proceder à atualização dos valores de $p^{(1)}$ e $p^{(2)}$, que envolve a mudança de dimensionalidade, usa-se o movimento com *reversible jump*.

Considera-se a passagem de um modelo de ordem p para um outro modelo de ordem p' e para que isso ocorra é necessário que se mantenha uma equação de equilíbrio detalhada que requer uma mapa determinístico invertível para estabelecer o ajuste de dimensões. Enciso-Mora *et al.* (2009) apresenta um procedimento para aumentar ou diminuir a ordem p numa unidade utilizando um mapeamento para dividir ou juntar os coeficientes autorregressivos para modelos $\text{INARMA}(p, q)$; tal procedimento é aqui estendido para os modelos $\text{SETINAR}(2; p^{(1)}, p^{(2)})$.

O modelo $\text{SETINAR}(2; p^{(1)}, p^{(2)})$ apresenta dois ramos em que as ordens podem ser iguais ou distintas. Determinar a ordem \mathbf{p} deste modelo é determinar o par $(p^{(1)}, p^{(2)})$ onde $p^{(1)}$ é a ordem do ramo superior e $p^{(2)}$ é a ordem do ramo inferior.

O primeiro passo é definir um dos ramos k para iniciar o movimento de uma ordem $p^{(k)}$ para uma ordem maior $p^{(k)} + 1$ ou para uma ordem inferior $p^{(k)} - 1$. Tem-se por objetivo manter $E[X_t]$ fixo em cada passo do procedimento pelo que, de acordo com Enciso-Mora *et al.* (2008), gera-se um novo $\alpha_i^{(k)'} \text{ tal que}$

$$\sum_{i=1}^{p'^{(k)}} \alpha_i^{(k)'} = \sum_{i=1}^{p^{(k)}} \alpha_i^{(k)}.$$

Definindo $\boldsymbol{\theta} := (\boldsymbol{\alpha}, \mathbf{p}, \boldsymbol{\lambda})$ e $\boldsymbol{\theta}' := (\boldsymbol{\alpha}', \mathbf{p}', \boldsymbol{\lambda}')$, a probabilidade de aceitação da mudança de ordem é dada por

$$A = 1 \wedge \underbrace{\frac{f(\mathbf{y}', \mathbf{z}', \boldsymbol{\theta}' | \mathbf{x})}{f(\mathbf{y}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{x})}}_I \times \underbrace{\frac{q(\mathbf{y}', \mathbf{z}', \boldsymbol{\theta}' \rightarrow \mathbf{y}, \mathbf{z}, \boldsymbol{\theta})}{q(\mathbf{y}, \mathbf{z}, \boldsymbol{\theta} \rightarrow \mathbf{y}', \mathbf{z}', \boldsymbol{\theta}')}}_{II} \times |J|,$$

onde J é o jacobiano das transformações a serem definidas posteriormente. De forma aleatória escolhe-se um dos ramos k para se propor uma nova ordem $p'^{(k)}$. No ramo escolhido, também de forma aleatória, escolhe-se um movimento de *nascimento* ou de *morte*. Green (1995) chama de nascimento quando se aumenta o espaço do parâmetro, e de morte quando se diminui.

Considere-se o movimento $p^{(k)} \rightarrow p'^{(k)} = p^{(k)} + 1$. Para se fazer a transformação $(\boldsymbol{\alpha}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}) \rightarrow (\boldsymbol{\alpha}', \mathbf{y}', \mathbf{z}, \boldsymbol{\lambda})$ tem-se o seguinte procedimento:

- (i) Gerar U de $U(0, 1)$;
- (ii) Gerar s de Uniforme discreta em $\{1, 2, \dots, p^{(k)}\}$;
- (iii) Para cada $i \in \{1, 2, \dots, p\}/s$, tomar $\alpha'_i{}^{(k)} = \alpha_i^{(k)}$;
- (iv) Considerar $\alpha'_s{}^{(k)} = U\alpha_s^{(k)}$ e $\alpha'_{p^{(k)}+1}{}^{(k)} = (1 - U)\alpha_s^{(k)}$;
- (v) Gerar $D_t \sim \text{Bin}(y_{t,s}, U)$ e a tomar $y'_{t,s} = D_t$ e $y'_{t,p^{(k)}+1} = y_{t,s} - D_t$ com $1 \leq t \leq n$. Os restantes permanecem fixos.

O movimento de um modelo $p^{(k)} \rightarrow p'^{(k)}$ é proposto com probabilidade dependente de s e U dada por

$$q(\mathbf{y}, \mathbf{z}, \boldsymbol{\theta} \rightarrow \mathbf{y}', \mathbf{z}', \boldsymbol{\theta}') = \left(\frac{1}{2}\right)^2 \times \frac{1}{p^{(k)}} \times 1 \times \prod_{t=1}^n \left(\frac{y_{t-s}}{y'_{t,s}} \right) U^{y_{t,s}} (1 - U)^{y_{t,s} - y'_{t,s}}.$$

O movimento contrário, o movimento de morte, acontece quando um dos coeficientes autorregressivos precisa ser eliminado. Para diminuir a ordem $p'^{(k)} + 1 \rightarrow p'^{(k)} = p^{(k)}$, gera-se somente um s de uma distribuição uniforme discreta entre 1 e $p^{(k)}$ e ajusta-se o termo $\alpha'_s = \alpha_s + \alpha_{p^{(k)}+1}$. O procedimento de juntar os termos y é

dado por $y'_{t,s} = y_{t,s} + y'_{t,p^{(k)}+1}$. A probabilidade proposta para o movimento do modelo $p^{(k)} + 1 \rightarrow p'^{(k)}$, depende neste caso somente da probabilidade de reduzir a ordem que é $\frac{1}{2}$ e a probabilidade de escolher o valor s , pelo o que é dada por

$$q(\mathbf{y}', \mathbf{z}', \theta' \rightarrow \mathbf{y}, \mathbf{z}, \theta) = \left(\frac{1}{2}\right)^2 \times \frac{1}{p^{(k)}}.$$

Assim, supondo que se está a aumentar a ordem $p^{(k)}$ do modelo no ramo k , o termo II da probabilidade de aceitação resume-se a

$$\frac{q(\mathbf{y}', \mathbf{z}', \theta' \rightarrow \mathbf{y}, \mathbf{z}, \theta)}{q(\mathbf{y}, \mathbf{z}, \theta \rightarrow \mathbf{y}', \mathbf{z}', \theta')} = \frac{1}{\prod_{t=1}^n \binom{y_{t-s}}{y'_{t,s}}} U^{y_{t,s}} (1-U)^{y_{t,s}-y'_{t,s}}.$$

O termo I para o cálculo da probabilidade de aceitação do movimento é dado a partir da razão entre as funções de verosimilhança conjuntas condicionadas a \mathbf{x} ,

$$\begin{aligned} \frac{f(\mathbf{y}', \mathbf{z}', \theta' | \mathbf{x})}{f(\mathbf{y}, \mathbf{z}, \theta | \mathbf{x})} &= \frac{\lambda^{(1)A^{(1)}} e^{-B^{(1)}\lambda^{(1)}} p'^{(1)} \pi(p'^{(1)}) \times \lambda^{(2)A^{(2)}} e^{-B^{(2)}\lambda^{(2)}} p'^{(2)} \pi(p'^{(2)})}{\lambda^{(1)A^{(1)}} e^{-B^{(1)}\lambda^{(1)}} p^{(1)} \pi(p^{(1)}) \times \lambda^{(2)A^{(2)}} e^{-B^{(2)}\lambda^{(2)}} p^{(2)} \pi(p^{(2)})} \\ &\times \prod_{t=1}^n \frac{\sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k)z^{(k)}} p'^{(k)}}{z^{(k)}!} \prod_{i=0}^{p'^{(k)}} \binom{x_{t-i}}{y'_{t,i}^{(k)}} \alpha_i'^{(k) y'_{t,i}^{(k)}} (1 - \alpha_i'^{(k)})^{x_{t-i} - y'_{t,i}^{(k)}}}{\sum_{k=1}^2 I_{t-d}^{(k)} \frac{e^{(-\lambda^{(k)})} \lambda^{(k)z^{(k)}} p^{(k)}}{z^{(k)}!} \prod_{i=0}^{p^{(k)}} \binom{x_{t-i}}{y_{t,i}^{(k)}} \alpha_i^{(k) y_{t,i}^{(k)}} (1 - \alpha_i^{(k)})^{x_{t-i} - y_{t,i}^{(k)}}}. \end{aligned}$$

Considerem-se separadamente possíveis movimentos e na perspectiva de se aumentar a ordem do regime $k = 1$ numa unidade, ou seja, $(p^{(1)}, p^{(2)}) \rightarrow (p'^{(1)}, p'^{(2)})$, com

$p'^{(1)} = p^{(1)} + 1$ e $p'^{(2)} = p^{(2)}$, tem-se

$$\frac{f(\mathbf{y}', \mathbf{z}', \theta' | \mathbf{x})}{f(\mathbf{y}, \mathbf{z}, \theta' | \mathbf{x})} = p^{(1)} \times \prod_{t=1}^n \frac{\sum_{k=1}^2 \left[I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z^{(1)}}}{z^{(1)}!} \binom{x_{t-s}}{y'_{t,s}^{(1)}} \alpha_s'^{(1) y'_{t,s}^{(1)}} (1 - \alpha_s'^{(1)})^{x_{t-s} - y'_{t,s}^{(1)}} \right]}{\sum_{k=1}^2 \left[I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z^{(1)}}}{z^{(1)}!} \binom{x_{t-s}}{y_{t,s}^{(1)}} \right]} \frac{\binom{x_{t-p'^{(1)}-1}}{y'_{t,p'^{(1)}+1}} \left(\alpha_{p'^{(1)}+1}'^{(1)} \right)^{y'_{t,p'^{(1)}+1}} \left(1 - \alpha_{p'^{(1)}+1}'^{(1)} \right)^{x_{t-p'^{(1)}-1} - y'_{t,p'^{(1)}+1}}}{\alpha_s^{(1) y_{t,s}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-s} - y_{t,s}^{(1)}}}.$$

Da mesma forma, aumentar a ordem do regime $k = 2$ numa unidade: $(p^{(1)}, p^{(2)}) \rightarrow (p'^{(1)}, p'^{(2)})$, com $p'^{(2)} = p^{(2)} + 1$ e $p'^{(1)} = p^{(1)}$, é dada por

$$\frac{f(\mathbf{y}', \mathbf{z}', \theta' | \mathbf{x})}{f(\mathbf{y}, \mathbf{z}, \theta' | \mathbf{x})} = p^{(2)} \times \prod_{t=1}^n \frac{\sum_{k=1}^2 \left[I_{t-d}^{(2)} \frac{e^{(-\lambda^{(2)})} \lambda^{(2)z^{(2)}}}{z^{(2)}!} \binom{x_{t-s}}{y'_{t,s}^{(2)}} \alpha_s'^{(2) y'_{t,s}^{(2)}} (1 - \alpha_s'^{(2)})^{x_{t-s} - y'_{t,s}^{(2)}} \right]}{\sum_{k=1}^2 \left[I_{t-d}^{(2)} \frac{e^{(-\lambda^{(2)})} \lambda^{(2)z^{(2)}}}{z^{(2)}!} \binom{x_{t-s}}{y_{t,s}^{(2)}} \right]} \frac{\binom{x_{t-p'^{(2)}-1}}{y'_{t,p'^{(2)}+1}} \left(\alpha_{p'^{(2)}+1}'^{(2)} \right)^{y'_{t,p'^{(2)}+1}} \left(1 - \alpha_{p'^{(2)}+1}'^{(2)} \right)^{x_{t-p'^{(2)}-1} - y'_{t,p'^{(2)}+1}}}{\alpha_s^{(2) y_{t,s}^{(2)}} (1 - \alpha_i^{(2)})^{x_{t-s} - y_{t,s}^{(2)}}}.$$

O último termo necessário para o cálculo da probabilidade de aceitação de um determinado modelo é o jacobiano que pode ser fatorizado em $n+1$ partes correspondendo à transformação $(\alpha^{(k)}, U) \rightarrow (\alpha_s^{(k)} U, (1-U)\alpha_s^{(k)})$ e $(y_{t,s}, D_t) \rightarrow (y'_{t,s}, y'_{t,p+1})$, $t = 1, \dots, n$ que conduz a $|J| = \alpha_s^{(k)} \times 1$. Tal traduz que o jacobiano é $|J| = \alpha_s^{(k)}$ se estivermos a alterar a ordem $p^{(k)}$, sendo $\alpha_s^{(k)}$ o valor de $\alpha^{(k)}$ correspondente ao parâmetro associado à saída s de uma distribuição uniforme $U\{1, 2, \dots, p^{(k)}\}$.

Este procedimento é tratado de uma forma similar, com pequenas adaptações, quando se pretende diminuir a ordem $p^{(k)}$ de uma unidade, em vez de aumentar.

2.4 Estudo de Simulação

O objetivo principal do estudo de simulação desta secção consiste em analisar o comportamento dos estimadores MQC, os de máxima verosimilhança, calculados pelo algoritmo EM e as estimativas bayesianas obtidas pelo (RJ)MCMC apresentados nas secções anteriores, variando as dimensões das amostras.

O estudo de simulação contempla quatro diferentes conjuntos de parâmetros $\theta = (\alpha^{(1)}, \alpha^{(2)}, \lambda^{(1)}, \lambda^{(2)}, d, r)$ apresentados na Tabela 2.1.

Tabela 2.1: Parâmetros dos modelos SETINAR(2; $p^{(1)}, p^{(2)}$)							
	$p^{(1)} = 1$		$p^{(2)} = 2$		$\lambda^{(1)}$	$\lambda^{(2)}$	r
	$\alpha_1^{(1)}$		$\alpha_1^{(2)}$	$\alpha_2^{(2)}$			
M_1	0.2		0.1	0.1	3	7	9
M_2	0.3		0.2	0.1	3	3	6
	$p^{(1)} = 2$		$p^{(2)} = 1$		$\lambda^{(1)}$	$\lambda^{(2)}$	r
	$\alpha_1^{(1)}$	$\alpha_2^{(1)}$	$\alpha_2^{(2)}$				
M_3	0.2	0.1	0.1		3	3	6
M_4	0.6	0.3	0.6		4	1	9

Os modelos M_1 e M_2 têm ordem 1 no primeiro regime e ordem 2 no segundo regime, enquanto que os modelos M_3 e M_4 apresentam ordens 2 e 1 respetivamente no primeiro e segundo regimes. Para cada combinação de parâmetros o valor do limiar r foi escolhido de forma que, pelo menos 50% das observações estivessem no primeiro regime e o parâmetro de atraso, d , foi considerado constante igual a 1.

A Figura 2.2 apresenta realizações dos quatro modelos considerados. As 3 linhas verticais ponteadas representadas em cada gráfico limitam superiormente as observações a serem consideradas para cada conjunto de dimensão $n = 50, 100, 200$. As situações (a) e (d), onde $\lambda^{(1)} \neq \lambda^{(2)}$, ilustram como o aumento do tamanho da

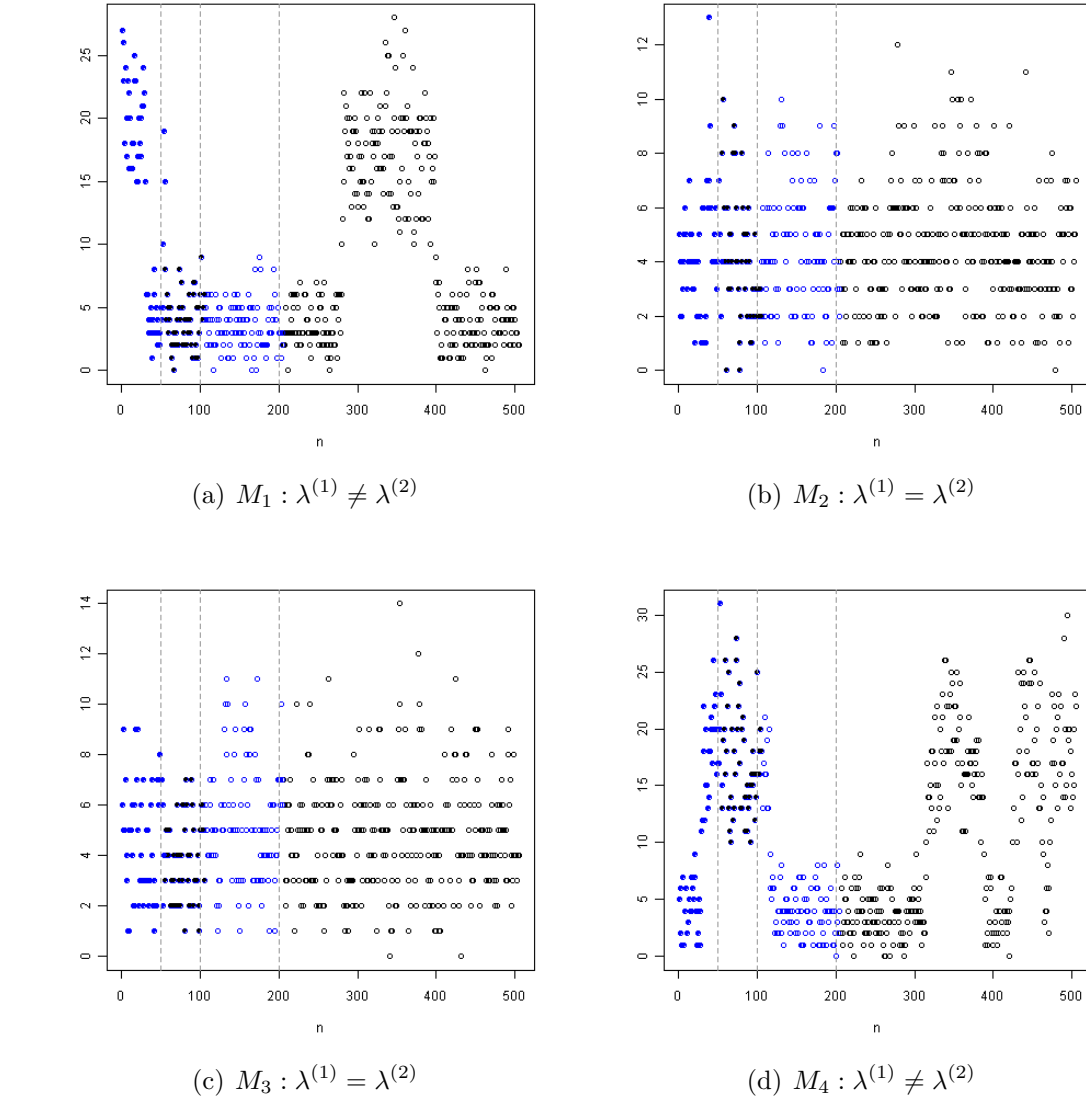


Figura 2.2: Realizações do modelo SETINAR

amostra incorpora informação relevante sobre o comportamento da série. Nos casos (b) e (c), onde $\lambda^{(1)} = \lambda^{(2)}$, a permanência no primeiro regime é superior à do segundo.

Para cada modelo simularam-se dados de dimensão 500 e consideraram-se as n primeiras observações para $n = 50, 100, 200, 500$; fizeram-se 500 réplicas. Com o objetivo de se proceder à comparação dos métodos de estimação anteriormente apresentados, apenas se consideraram as realizações cujos valores estimados para os parâmetros fossem admissíveis.

Para calcular as estimativas bayesianas, usando os algoritmos MCMC ou RJMCMC,

correspondentes a $p^{(1)}$ e $p^{(2)}$ serem conhecidas ou não, respetivamente, para cada conjunto de dados de dimensão $n = 50, 100, 200, 500$ simulou-se uma única cadeia de dimensão 10000, com um período de *burn-in* de 21000 e desfasamento 20 entre as observações selecionadas. Foram consideradas as distribuições *a priori* não informativas, supondo $A \approx 0$ e $B \approx 0$, de acordo com a especificação feita na Secção 2.3.3.

A apresentação e discussão dos resultados está dividida em duas partes. Na primeira parte faz-se uma análise comparativa dos 3 métodos de estimação apresentados e a segunda parte debruça-se sobre a questão da seleção de modelos.

2.4.1 Análise dos Resultados de Estimação

As médias amostrais obtidas para os parâmetros estimados $\hat{\theta}$ dos modelos M_1 e M_2 são apresentadas na Tabela 2.2 com os respetivos erros quadráticos médios (EQM) em parêntesis e os correspondentes viés são apresentados nas Figuras 2.3 e 2.4. Ao fazer a análise dos resultados pode-se dizer que, globalmente as estimativas obtidas pelos métodos MQC e MVC - referidas por EM - têm pior desempenho e apresentam uma maior dispersão do que as estimativas bayesianas - referidas por MCMC. Para todos os parâmetros, tanto para as amostras pequenas ($n = 50$) como para as amostras grandes ($n = 500$), a metodologia bayesiana MCMC apresenta menor viés nas estimativas.

Para o modelo M_1 , o algoritmo MCMC conduz a estimativas bayesianas com menor erro quadrático médio para todos os valores estimados de $\alpha_1^{(2)}$ e $\alpha_2^{(2)}$, no entanto as boas propriedades assintóticas dos estimadores de mínimos quadrados e de máxima verosimilhança são notórias para estimar o parâmetros $\lambda_1^{(1)}$ e $\alpha_1^{(2)}$, com $n = 500$. Neste modelo, deve-se ainda destacar a dificuldade que qualquer uma das metodologias apresenta para estimar o parâmetro $\lambda^{(2)}$, sobrestimando-o (com exceção da estimativa bayesiana com $n = 500$) e apresentando um grande enviesamento. Considerando o modelo M_2 as estimativas bayesianas continuam a ter globalmente um melhor desempenho, quer em termos do enviesamento, quer em termos do seu EQM; apenas apresenta um comportamento ligeiramente menos bom para estimar $\lambda^{(1)}$ para

$n \neq 50$ e pior para estimar $\alpha_2^{(2)}$, com $n = 200$. É de referir que, apesar de os valores médios das inovações serem iguais entre si, isto é, $\lambda^{(k)} = 3$, para $k = 1, 2$, o enviesamento das estimativas para estimar o parâmetro da inovação do 2º regime é bastante superior. As médias amostrais obtidas para os parâmetros estimados $\hat{\theta}$ dos modelos M_1 e M_2 são apresentadas na Tabela 2.2 com os respetivos erros quadráticos médios (EQM) em parêntesis e os correspondentes viés são apresentados nas Figuras 2.3 e 2.4. Seguindo um procedimento análogo ao anteriormente feito para os modelos M_1 e M_2 os resultados obtidos para os modelos M_3 e M_4 estão apresentados na Tabela 2.3 e o viés das estimativas estão apresentados nas Figuras 2.5 e 2.6. Uma característica que diferencia os modelos M_3 e M_4 é a escolha dos valores dos parâmetros autorregressivos α de cada um dos regimes para gerar as amostras. No modelo M_3 os valores escolhidos são pequenos, $\alpha^{(1)} = (0.2, 0.1)$ e $\alpha^{(2)} = (0.1)$, enquanto que para o modelo M_4 optou-se por valores maiores, $\alpha^{(1)} = (0.6, 0.3)$ e $\alpha^{(2)} = (0.6)$, mas que satisfizessem as condições de estacionariedade impostas, $\sum \alpha^{(k)} < 1$. As estimativas obtidas para o parâmetro $\alpha^{(1)}$, nos modelos M_3 e M_4 , apresentam comportamento distinto. Quando os verdadeiros valores de $\alpha^{(1)}$ são pequenos (M_3) o viés é menor do que quando o verdadeiro valor do parâmetro é grande (M_4). O parâmetro $\alpha^{(2)}$ possui comportamento contrário: quando o valor real do parâmetro é pequeno (M_3) o enviesamento das estimativas é maior e o parâmetro é sobrestimado, enquanto que, quando $\alpha^{(2)}$ é grande (M_4) o viés é menor e o parâmetro é subestimado. Os modelos M_1 e M_4 possuem $\lambda^{(1)} \neq \lambda^{(2)}$ e em ambos os casos $\hat{\lambda}$ para o ramo de ordem 2 apresentou valores sobrestimados com grande viés.

Globalmente verifica-se que, em cada regime, o mau(bom) desempenho de qualquer método para estimar os parâmetros autorregressivos está associado ao mau(bom) desempenho das estimativas dos parâmetros dos processos de inovação. O enviesamento das estimativas dos parâmetros diminui à medida que a dimensão das realizações aumenta, o que está de acordo com as propriedades dos estimadores de serem assintoticamente centrados e consistentes.

Tabela 2.2: Média amostral e EQM (em parêntesis) das estimativas dos parâmetros para os modelos M_1 e M_2

Modelo		n															
		50				100				200				500			
$p^{(1)} = 1$		MQC	EM	MCMC		MQC	EM	MCMC		MQC	EM	MCMC		MQC	EM	MCMC	
$p^{(2)} = 2$	θ_i																
M_1	$\alpha_1^{(1)}$	0.2777 (0.0385)	0.2580 (0.0333)	0.2545 (0.0140)		0.2394 (0.0216)	0.2233 (0.0186)	0.2372 (0.0105)		0.2114 (0.0086)	0.2010 (0.0081)	0.2012 (0.0018)		0.2001 (0.0026)	0.2008 (0.0027)	0.2184 (0.0050)	
	0.2																
	$\alpha_1^{(2)}$	0.1635 (0.0203)	0.1664 (0.0219)	0.1804 (0.0084)		0.1598 (0.0160)	0.1527 (0.0081)	0.1554 (0.0040)		0.1340 (0.0138)	0.1196 (0.0039)	0.1280 (0.0010)		0.1243 (0.0094)	0.1089 (0.0070)	0.1484 (0.0029)	
	0.1																
	$\alpha_2^{(2)}$	0.1592 (0.0182)	0.1559 (0.0181)	0.1235 (0.0015)		0.1601 (0.0170)	0.1332 (0.0091)	0.1043 (0.0000)		0.1479 (0.0123)	0.1160 (0.0032)	0.1075 (0.0000)		0.1133 (0.0071)	0.1126 (0.0057)	0.1017 (0.0000)	
	0.1																
	$\lambda^{(1)}$	2.9000 (0.7181)	2.9080 (0.6743)	2.2980 (0.5060)		2.9760 (0.3571)	2.9660 (0.3127)	2.6700 (0.1934)		3.0120 (0.1462)	2.9910 (0.1201)	2.9960 (0.0426)		3.0110 (0.0452)	2.9960 (0.0426)	3.2190 (0.3523)	
	3																
	$\lambda^{(2)}$	11.7800 (33.0469)	8.2548 (7.8677)	8.5240 (9.1686)		11.7473 (31.0454)	7.8716 (6.9901)	8.1860 (4.9897)		12.0790 (32.5985)	8.1030 (6.6347)	7.9630 (2.9291)		12.7530 (37.9565)	8.0249 (3.9264)	6.6980 (1.1654)	
	7																
p			0.7180				0.7980				0.8780				0.8740		
M_2	$\alpha_1^{(1)}$	0.3053 (0.0289)	0.3146 (0.0305)	0.3042 (0.0209)		0.3070 (0.0173)	0.3094 (0.0161)	0.3026 (0.0133)		0.3035 (0.0094)	0.3057 (0.0089)	0.2916 (0.0094)		0.2971 (0.0052)	0.3026 (0.0058)	0.3014 (0.0009)	
	0.3																
	$\alpha_1^{(2)}$	0.2715 (0.0420)	0.2833 (0.0405)	0.2625 (0.0286)		0.2707 (0.0358)	0.2604 (0.0316)	0.2435 (0.0180)		0.2755 (0.0327)	0.2517 (0.0187)	0.2367 (0.0087)		0.2494 (0.0292)	0.2497 (0.0287)	0.1985 (0.0020)	
	0.2																
	$\alpha_2^{(2)}$	0.2070 (0.0364)	0.2010 (0.0354)	0.2058 (0.0281)		0.1964 (0.0260)	0.1862 (0.0260)	0.2012 (0.0193)		0.1639 (0.0153)	0.1594 (0.0070)	0.1932 (0.0148)		0.1463 (0.0226)	0.1090 (0.0078)	0.1367 (0.0024)	
	0.1																
	$\lambda^{(1)}$	3.0940 (0.4613)	3.0572 (0.4888)	2.8150 (0.2186)		3.0220 (0.2550)	2.9840 (0.2538)	2.9040 (0.2351)		3.0200 (0.1387)	3.0190 (0.1468)	2.8990 (0.1348)		2.6900 (0.1448)	2.7800 (0.0921)	2.9040 (0.1145)	
	3																
	$\lambda^{(2)}$	2.8615 (2.7580)	2.8100 (2.8295)	2.8730 (1.8502)		2.9408 (2.3006)	3.0590 (1.2394)	2.6480 (1.0043)		3.0811 (1.8829)	3.0257 (1.6577)	2.8990 (0.8764)		2.8790 (3.1785)	3.0050 (2.1259)	2.8920 (1.4211)	
	3																
	p		0.7852				0.8039				0.8146				0.8226		

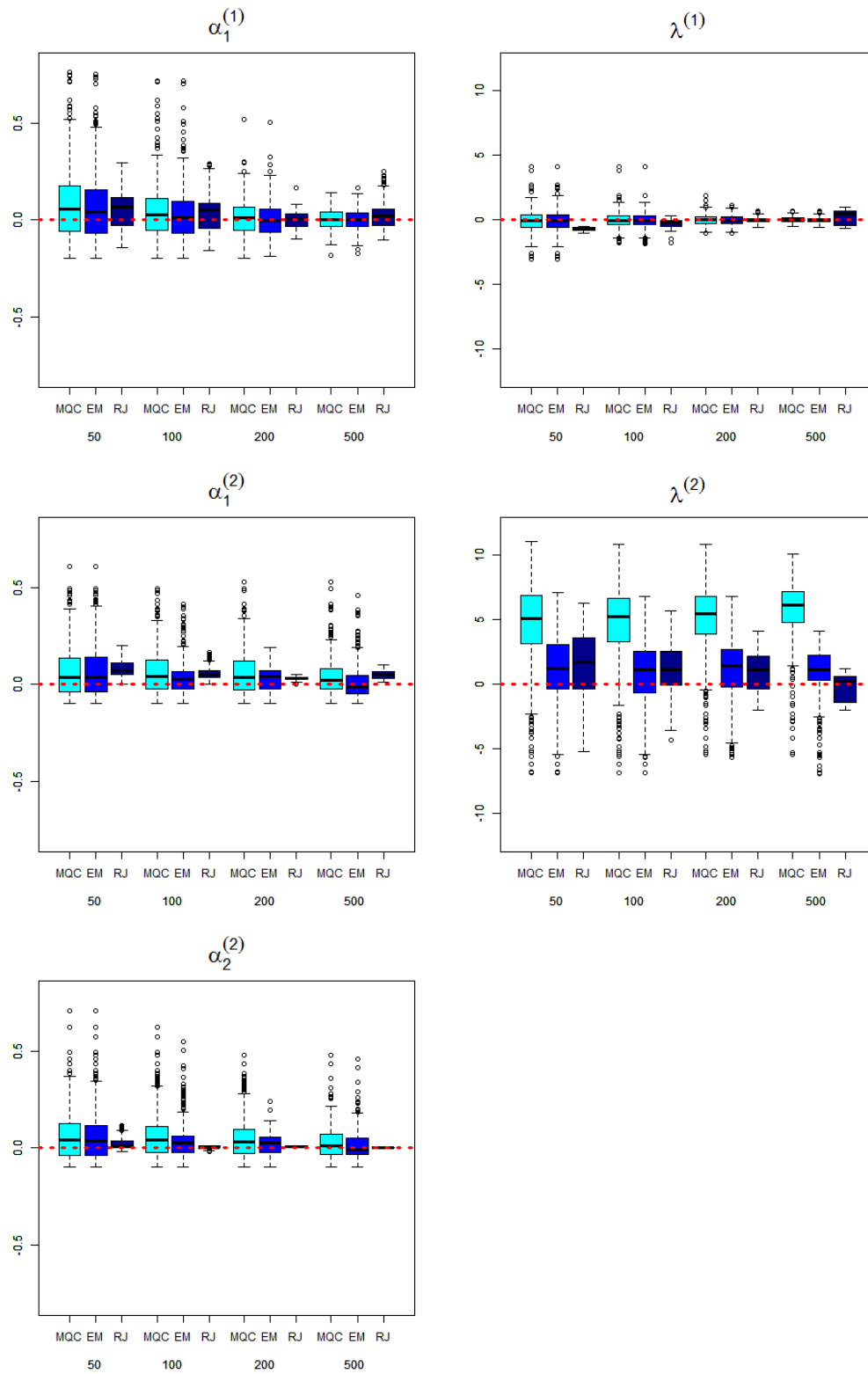


Figura 2.3: Viés das estimativas de $\theta = (0.2; 0.1, 0.1; 3; 7)$ do modelo M_1 .

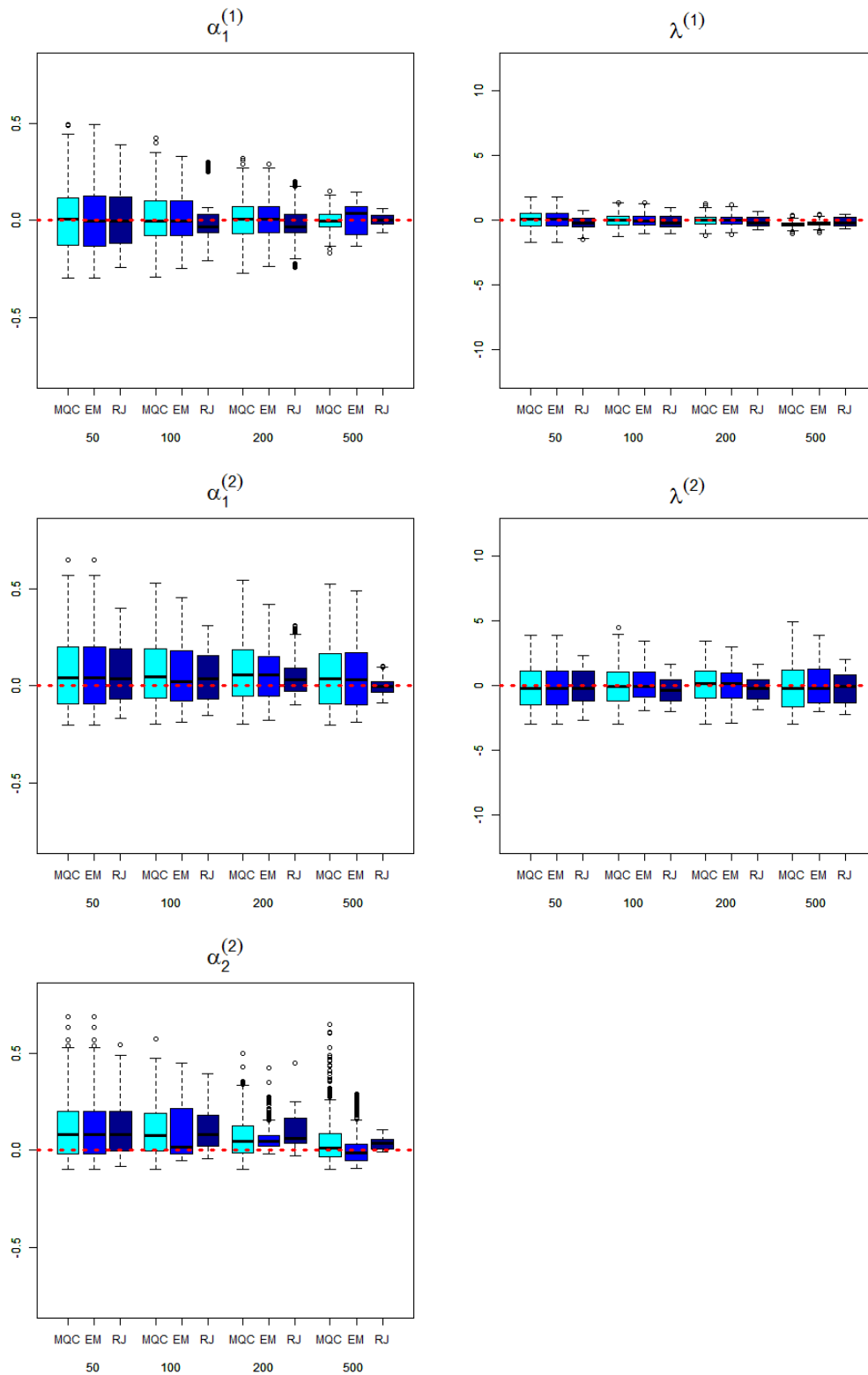


Figura 2.4: Viés das estimativas de $\theta = (0.3; 0.2, 0.1; 3; 3)$ do modelo M_2 .

Tabela 2.3: Média amostral e EQM (em parêntesis) das estimativas dos parâmetros para os modelos M_3 e M_4

Modelo	θ_i	n					
		50		100		200	
		MQC	EM	MC	MC	MQC	EM
M_3	$p^{(1)} = 2$						
	$\alpha_1^{(1)}$	0.2826	0.2601	0.2355	0.2245	0.2812	0.2481
	0.2	(0.0325)	(0.0171)	(0.0131)	(0.0092)	(0.0148)	(0.0116)
	$\alpha_2^{(1)}$	0.1040	0.1176	0.1351	0.1183	0.0643	0.1181
	0.1	(0.0075)	(0.0103)	(0.0045)	(0.0036)	(0.0032)	(0.0084)
	$\alpha_1^{(2)}$	0.3486	0.2425	0.2335	0.2181	0.3260	0.1948
	0.1	(0.1088)	(0.0554)	(0.0421)	(0.0308)	(0.0865)	(0.0224)
	$\lambda^{(1)}$	2.7053	2.7607	2.7580	2.8040	2.8360	2.8970
	3	(0.5372)	(0.4931)	(0.4619)	(0.4520)	(0.1751)	(0.1757)
	$\lambda^{(2)}$	3.4530	3.1444	3.466	3.1830	3.6508	3.0450
M_4	3	(3.5501)	(1.9806)	(1.3756)	(0.9143)	(2.8272)	(0.6573)
	p		0.7967	0.8116			0.8194
	$\alpha_1^{(1)}$	0.5042	0.4940	0.5188	0.6049	0.5515	0.5793
	0.6	(0.0840)	(0.0424)	(0.0307)	(0.0093)	(0.0592)	(0.0096)
	$\alpha_2^{(1)}$	0.4160	0.3918	0.3731	0.3348	0.3487	0.3458
	0.3	(0.0840)	(0.0424)	(0.0307)	(0.0110)	(0.0281)	(0.0250)
	$\alpha_1^{(2)}$	0.5572	0.5275	0.5871	0.5758	0.5832	0.5860
	0.6	(0.0086)	(0.0135)	(0.0066)	(0.0052)	(0.0025)	(0.0024)
	$\lambda^{(1)}$	7.3633	6.3210	5.343	5.3820	7.7710	6.4990
	4	(20.4992)	(8.9463)	(5.2146)	(5.0726)	(18.7687)	(7.1932)
M_4	$\lambda^{(2)}$	1.5229	1.4680	1.3198	1.1663	1.1969	1.1057
	1	(1.3423)	(1.0450)	(0.7167)	(0.4250)	(0.3984)	(0.3275)
	p		0.7009	0.7156			0.7688
	$\alpha_1^{(1)}$	0.5761	0.5932	0.5919	0.5919	0.5761	0.5932
	0.6	(0.0349)	(0.0237)	(0.0102)	(0.0102)	(0.0349)	(0.0237)
	$\alpha_2^{(1)}$	0.3157	0.3151	0.3336	0.3336	0.3157	0.3151
	0.3	(0.0131)	(0.0127)	(0.0065)	(0.0065)	(0.0131)	(0.0127)
	$\alpha_1^{(2)}$	0.5956	0.5948	0.5765	0.5765	0.5956	0.5948
	0.6	(0.0011)	(0.0011)	(0.0024)	(0.0024)	(0.0011)	(0.0011)
	$\lambda^{(1)}$	7.9620	5.7090	5.5540	5.5540	7.9620	5.7090
M_4	4	(17.9331)	(3.5146)	(5.1674)	(5.1674)	(17.9331)	(3.5146)
	$\lambda^{(2)}$	1.0541	1.0540	1.1468	1.1468	1.0541	1.0540
	1	(0.1757)	(0.1771)	(0.3236)	(0.3236)	(0.1757)	(0.1771)
	p		0.7726				0.7726

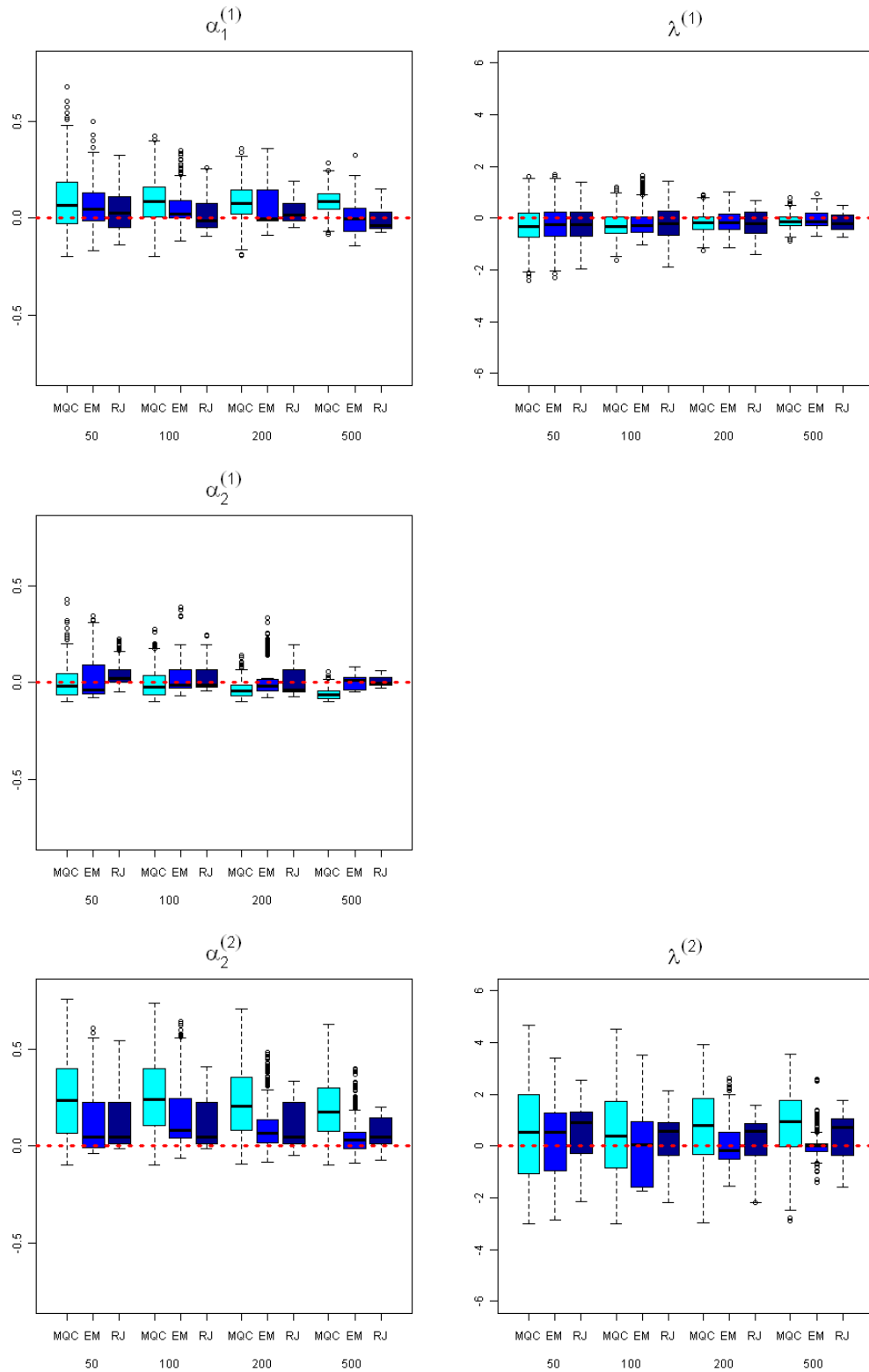


Figura 2.5: Viés das estimativas de $\theta = (0.2, 0.1; 0.1; 3; 3)$ do modelo M_3 .

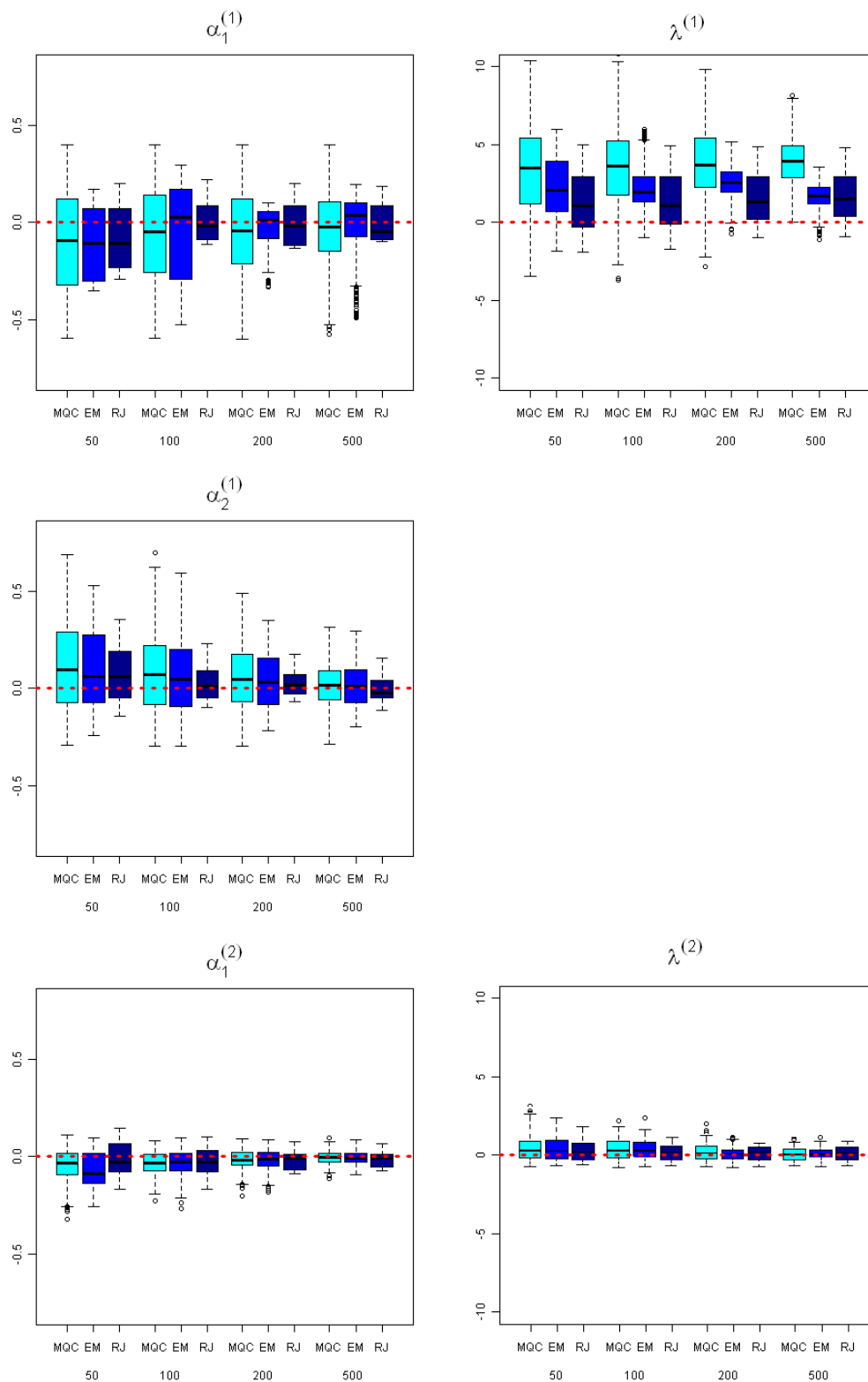


Figura 2.6: Viés das estimativas de $\theta = (0.6, 0.3; 0.6; 4; 1)$ do modelo M_4 .

2.4.2 Seleção de Modelos

Com vista a fazer a seleção de modelos segundo a metodologia clássica, os modelos podem ser comparados usando critérios de informação baseados na verosimilhança, em particular o critério de informação de Akaike (Akaike, 1973) designado por AIC (*Akaike Information Criterion*) ou o critério de informação bayesiano (*Bayesian Information Criterion*), introduzido por Schwarz (1978) e designado por BIC. Alternativamente existe uma série de aproximações bayesianas para fazer o ajustamento e seleção de modelos, destacando-se o DIC (*Deviance Information Criterion*) e os Fatores de Bayes (Spiegelhalter *et al.*, 2002). No entanto o algoritmo de Monte Carlo via cadeias de Markov com saltos reversíveis - RJMCMC - introduzido por Green (1995), permite tratar simultaneamente a questão da seleção do modelo e da estimação dos parâmetros.

Nesta secção irão ser usados os critérios BIC, DIC e o algoritmo RJMCMC. Assim, o modelo selecionado será o que apresentar menor valor de

$$BIC(x) := -2 \ln L(\hat{\theta}_i | x) + p \ln n,$$

onde $\hat{\theta}$ são as estimativas de máxima verosimilhança, p o número de parâmetros do modelo e n a dimensão da amostra, ou de

$$DIC(x) := 2E_{\theta|x}[D(\theta)] - D(\bar{\theta}),$$

onde $\bar{\theta} = E_{\theta|x}(\theta)$ e $D(\theta) = -2 \ln f(x|\theta)$, segundo os critérios BIC e DIC, respetivamente. No caso de se utilizar o algoritmo RJMCMC escolhe-se o modelo para o qual a cadeia convergiu.

Nas Tabelas 2.4 e 2.5 estão apresentados os valores do BIC e DIC obtidos. Os modelos corretos estão destacados a negrito e os valores marcados com * indicam a combinação modelo-dimensão da amostra que possui os menores valores de BIC e

DIC, ou seja, a combinação considerada correta para o modelo em questão. Pode-se verificar que para todos os 4 modelos, qualquer um dos critérios escolheu o modelo correto.

Tabela 2.4: Seleção de modelos pelo critério BIC.

Modelo	n	$(p^{(1)}, p^{(2)})$			
		(1,1)	(1,2)	(2,1)	(2,2)
M_1	50	1252.16	1154.21	1553.12	1354.21
	100	1298.10	1080.61	1194.17	1124.02
	200	1325.14	851.02	1258.21	1111.87
	500	1270.00	832.22*	1227.29	1102.13
M_2	50	1785.12	1244.77	1985.12	1752.00
	100	1244.77	780.16	1744.77	1542.01
	200	1231.27	681.01	1531.27	1354.20
	500	1222.99	675.23*	1422.99	1309.22
M_3	50	1012.12	1985.12	1512.12	1655.00
	100	999.54	1744.77	1004.98	1584.01
	200	1000.12	1531.27	1000.12	1901.00
	500	982.51	1422.99	981.50*	1458.96
M_4	50	1458.03	1122.12	1098.36	1374.23
	100	1268.00	1164.03	909.01	1325.64
	200	1321.02	1002.44	689.01	962.27
	500	1265.19	974.73	688.09*	974.74

Para fazer uma análise do desempenho do RJMCMC, repetiu-se o algoritmo 500 vezes e calculou-se a percentagem de escolha de ordem correta; os resultados podem ser observados na Tabela 2.6 e as Figuras 2.7 a 2.10 indicam o número de vezes que cada modelo foi escolhido.

Quando a dimensão da amostra aumenta ela torna-se mais informativa melhorando o desempenho do algoritmo tal como constatado por Enciso-Mora *et al.* (2008).

Na Tabela 2.7 apresentam-se os resultados obtidos usando as 3 abordagens.

Tabela 2.5: Seleção de modelos pelo critério DIC.

Modelo	n	$(p^{(1)}, p^{(2)})$			
		(1,1)	(1,2)	(2,1)	(2,2)
M_1	50	691	641	711	699
	100	690	638	725	698
	200	685	627	687	688
	500	674	624*	667	661
M_2	50	715	712	717	865
	100	718	709	718	882
	200	801	702	708	879
	500	798	698*	708	867
M_3	50	709	698	689	739
	100	722	699	612	716
	200	710	702	591	766
	500	716	702	589*	763
M_4	50	839	865	849	1045
	100	841	857	826	1030
	200	840	842	825	1033
	500	856	869	822*	1041

Tabela 2.6: Seleção de modelos usando o algoritmo RJMCMC

Modelo	n	$(p^{(1)}, p^{(2)})$			
		(1,1)	(1,2)	(2,1)	(2,2)
		%	%	%	%
M_1	50	28	38	151	19
	100	26	43	14	17
	200	20	48	14	18
	500	15	49*	17	19
M_2	50	30	35	25	10
	100	28	40	20	12
	200	18	50	21	11
	500	17	51*	21	10
M_3	50	24	30	32	14
	100	14	26	44	16
	200	22	20	51	7
	500	19	22	56*	3
M_4	50	33	28	30	9
	100	30	29	31	12
	200	30	30	31	11
	500	28	28	34*	10

Tabela 2.7: Seleção de modelos usando BIC, DIC e RJMCMC

Modelo	n	(1,1)			(1,2)			(2,1)			(2,2)		
		BIC	%	DIC	BIC	%	DIC	BIC	%	DIC	BIC	%	DIC
M_1	50	1252.16	28	691	1154.21	38	641	1553.12	15	711	1354.21	19	699
	100	1298.10	26	690	1080.61	43	638	1194.17	14	725	1124.02	17	698
	200	1325.14	20	685	851.02	48	627	1258.21	14	687	1111.87	18	688
	500	1270.00	15	674	832.22*	49	624	1227.29	17	667	1102.13	19	661
M_2	50	1785.12	30	715	1244.77	35	712	1985.12	25	717	1752.00	10	865
	100	1244.77	28	718	780.16	40	709	1744.77	20	718	1542.01	12	882
	200	1231.27	18	801	681.01	50	702	1531.27	21	708	1354.20	11	879
	500	1222.99	17	798	675.23*	52	698	1422.99	21	708	1309.22	10	867
M_3	50	1012.12	24	709	1985.12	30	698	1512.12	32	689	1655.00	14	739
	100	999.54	14	722	1744.77	26	699	1004.98	44	612	1584.01	16	716
	200	1000.12	22	710	1531.27	20	702	1000.12	51	591	1901.00	7	766
	500	982.51	19	716	1422.99	22	702	981.50*	56	589	1458.96	3	763
M_4	50	1458.03	33	839	1122.12	28	865	1098.36	30	849	1374.23	9	1045
	100	1268.00	30	841	1164.03	29	857	909.01	31	826	1325.64	12	1030
	200	1321.02	30	840	1002.44	30	842	689.01	31	825	962.27	11	1033
	500	1265.19	28	856	974.73	28	869	688.09*	34	822	974.74	10	1041

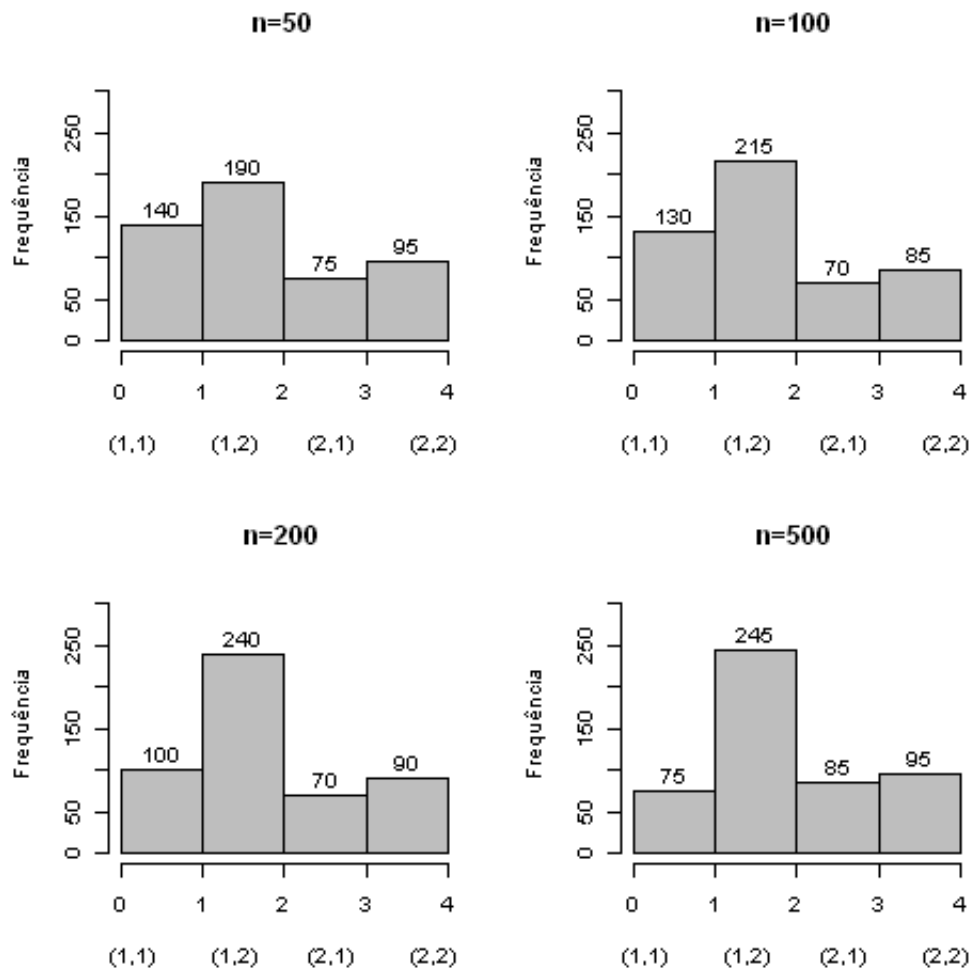


Figura 2.7: Modelo M_1 .

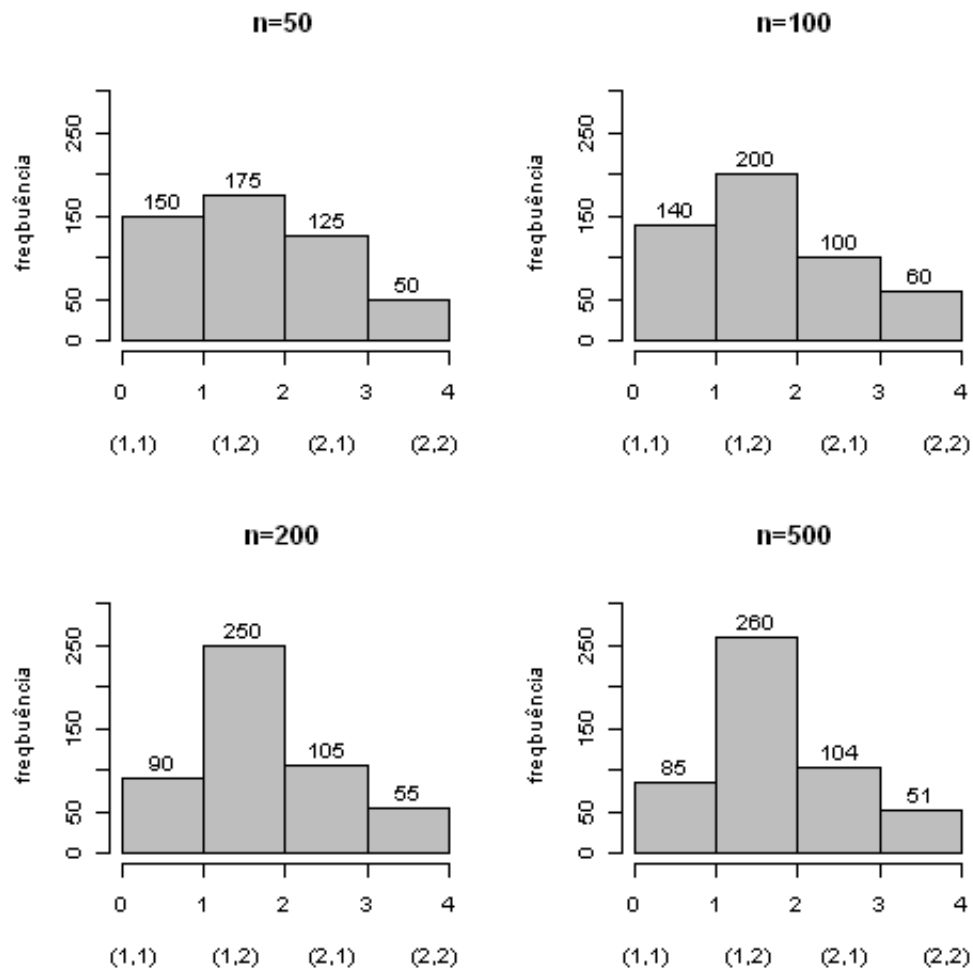


Figura 2.8: Modelo M_2 .

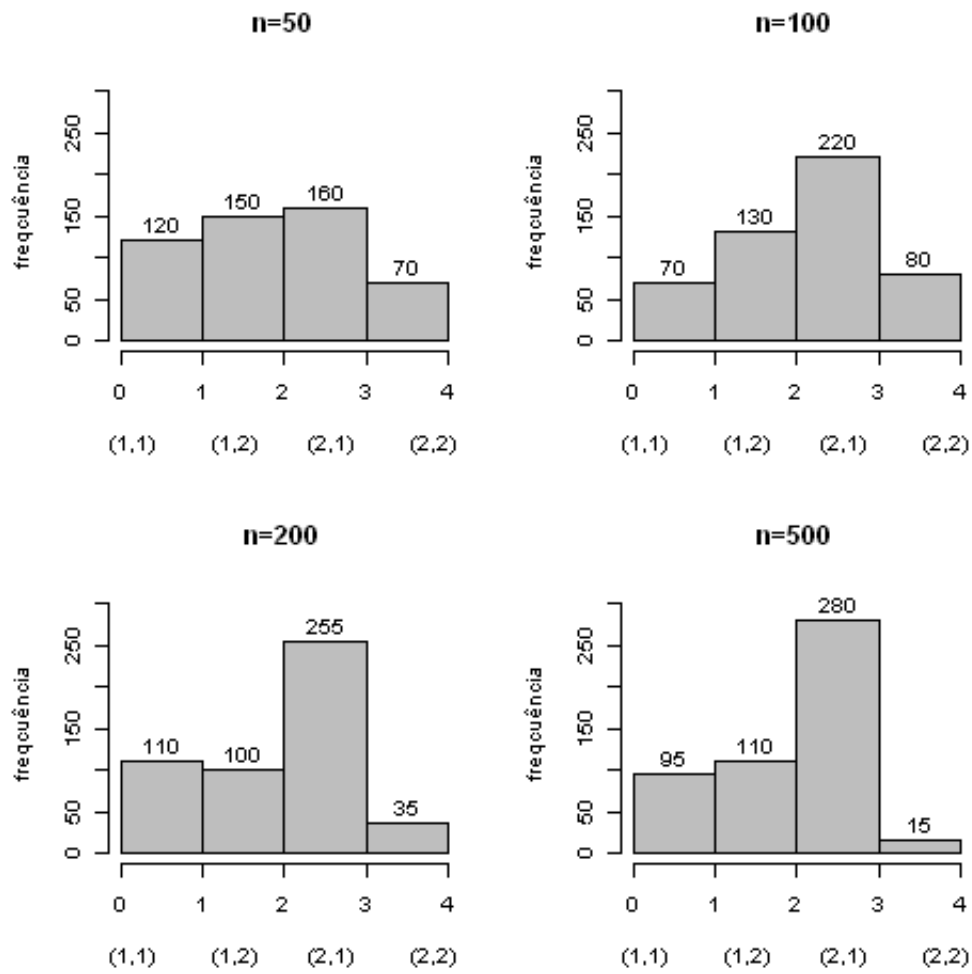


Figura 2.9: Modelo M_3 .

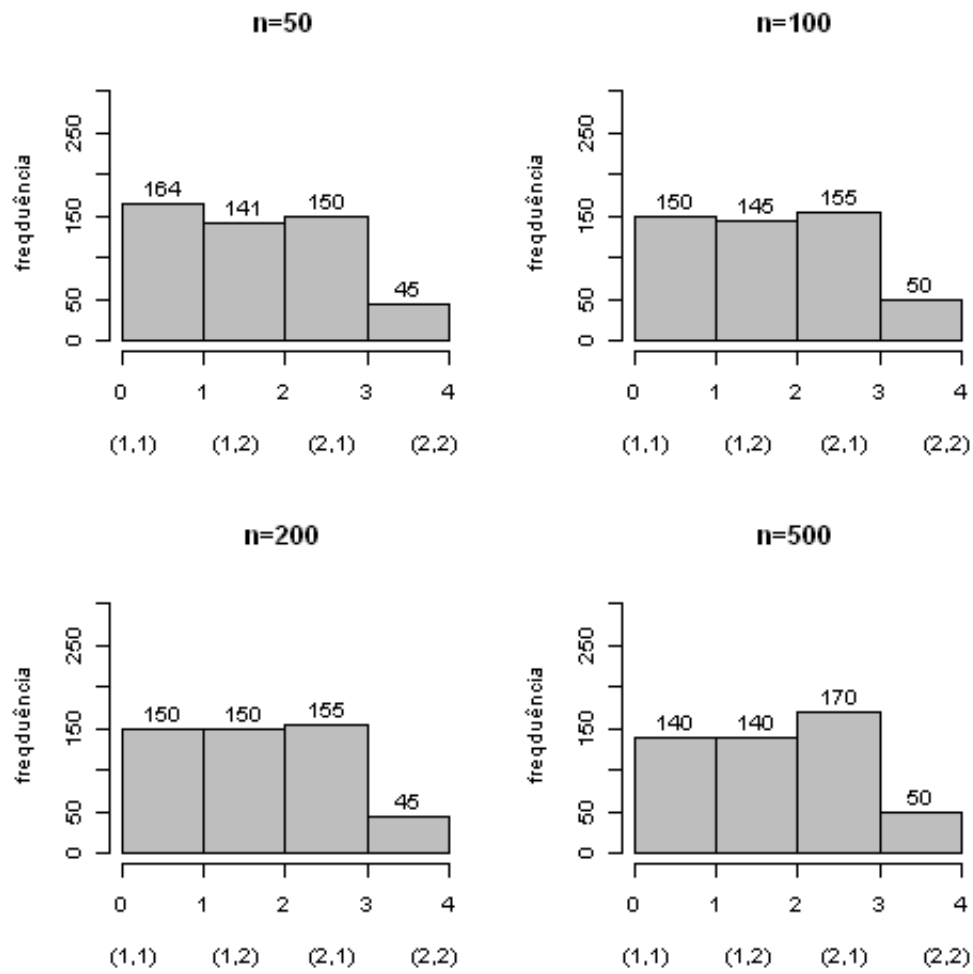


Figura 2.10: Modelo M_4 .

2.5 Aplicação a Dados Reais

Nesta secção usa-se dois conjuntos de dados relativos a manchas solares para ilustrar o modelo introduzido na equação (2.1).

Uma mancha solar é uma região na superfície solar, fotosfera, que está temporariamente mais fria (2000K) e mais escura por contraste comparada as regiões que a circundam (5700K). A região denominada *mancha* emite menos energia por possuir um campo magnético mais intenso o qual bloqueia parcialmente a energia transmitida por convecção. Tais regiões apresentam campos magnéticos milhares de vezes mais forte que o campo magnético da terra. As manchas mudam a cada momento em forma e intensidade e a sua duração pode ser de poucos dias até meses, tendo em média a duração de 2 semanas. Em geral surgem em grupos de 2 conjuntos de manchas, onde um conjunto formará o campo magnético positivo e o outro conjunto o campo magnético negativo.

A região central mais escura das manchas solares, chamada de umbra, apresenta o campo magnético mais forte. A parte mais clara, chamada penumbra, apresenta aproximadamente 2.5 vezes o tamanho da umbra e seu campo magnético é mais fraco. Sabe-se também que nem todas as manchas apresentam penumbra.

A análise das séries temporais de observação de manchas solares mostram um ciclo regular médio da atividade solar de aproximadamente 11 anos com variação da sua forma e amplitude. Rudolf Wolf em 1848 (ver Izenman, 1983) introduziu o número relativo de manchas solares como medida da atividade de manchas solares. As séries temporais registradas por Wolf foram compiladas após sua morte por seu assistente Wolfer e também são conhecidas como *séries de manchas solares de Wolfer*. Hoyt & Schatten (1998) fazem uso do número de grupos de manchas observadas ao invés de manchas individuais.

Box & Jenkins (1976) usaram o número de manchas solares de Wolfer para demonstrar suas técnicas de previsão na análise de séries temporais e verificaram que um modelo ótimo para os dados de 1770-1869 seria um modelo autorregressivo de 2º

ou 3º ordem. Wei (1989) ampliou o estudo feito por Box & Jenkins (1976) e sugere que a periodicidade do número de manchas solares é de 11 anos. Tong (1983) fez uso de 280 observações do período de 1700-1979 analisando a série temporal sob a perspectiva de um modelo *threshold*.

Stramer & Lin (2002) usaram os mesmos dados de Tong (1983) numa perspectiva bayesiana e o algoritmo RJMCMC. Campbell (2004) fez uso dos dados médios de manchas solares (1749-1984) para ajustar um modelo SETAR com uma cadeia de Markov com saltos reversíveis. Silva *et al.* (2005) consideraram o número total de manchas solares semanais do período de 2 anos (1990-1991) para ajustarem um modelo autorregressivo de valores inteiros. Monteiro *et al.* (2008) desenvolveram um sistema de alarme ótimo para o modelo DSINAR(1) e exemplificaram usando dados diários de manchas solares (dez/1998-abr/1999).

A Figura 2.11 representa um grupo de manchas solares na região ativa do sol identificada por *Active Region 10030* observada em 15/7/2002. A imagem foi obtida pelo telescópio solar Swedish 1-m¹ pertence a Academia Real de Ciências da Suécia.²

2.5.1 Aplicação 1

Os dados aqui utilizados foram retirados de Box *et al.* (2008) e são relativos à série clássica do número anual de manchas solares de Wolfer entre os anos de 1770 e 1869.

Para a série de Wolfer usou-se como limiar $r = 34$ e como desfasamento $d = 11$. A escolha do valor do limiar foi baseada em valores discutidos por Campbell (2004). O limiar foi escolhido baseado no ciclo aproximado de 11 anos das manchas solares sugerido em diversos trabalhos. Com vista à seleção do modelo, consideraram-se todas as combinações dos modelos SETINAR(2; $p^{(1)}, p^{(2)}$), $1 \leq p^{(k)} \leq 2$, $k = 1, 2$ e

¹Swedish 1-m Solar Telescope ou SST - é um telescópio solar de refração situado no Observatório Roque de los Muchachos em La Palma nas Ilhas Canárias.

²Em inglês - *Royal Swedish Academy of Sciences*.

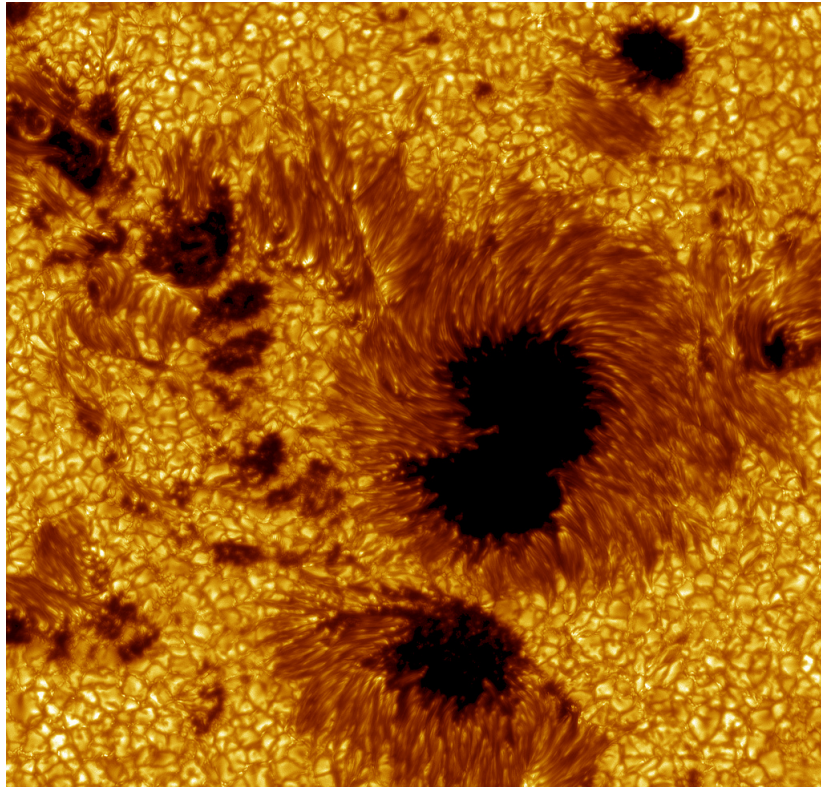


Figura 2.11: Mancha solar - 15/7/2002 - Royal Swedish Academy of Sciences.

aplicaram-se os critérios BIC, DIC e o algoritmo RJMCMC. De acordo com os resultados apresentados na Tabela 2.8 foi selecionado o modelo $\text{SETINAR}(2; 2, 1)$. As estimativas dos parâmetros encontram-se na Tabela 2.9.

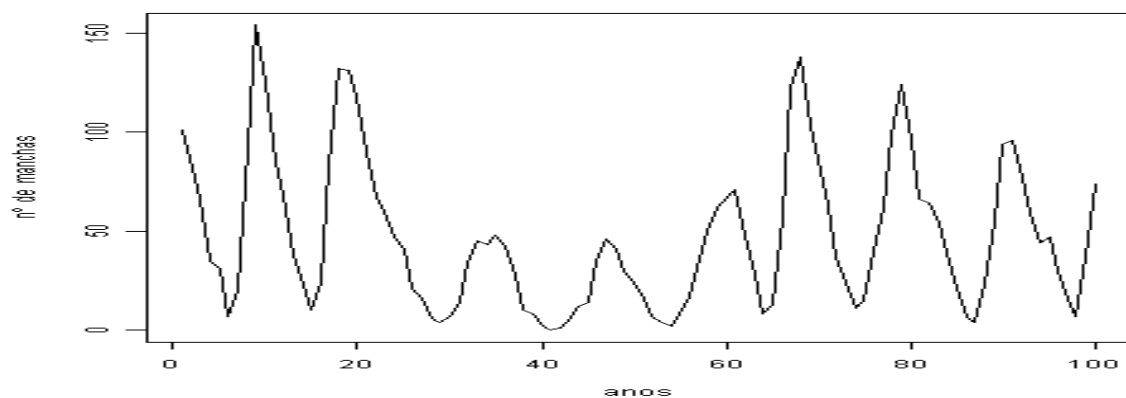


Figura 2.12: Série temporal do número anual de manchas solares de Wolfer ano período 1770-1869.

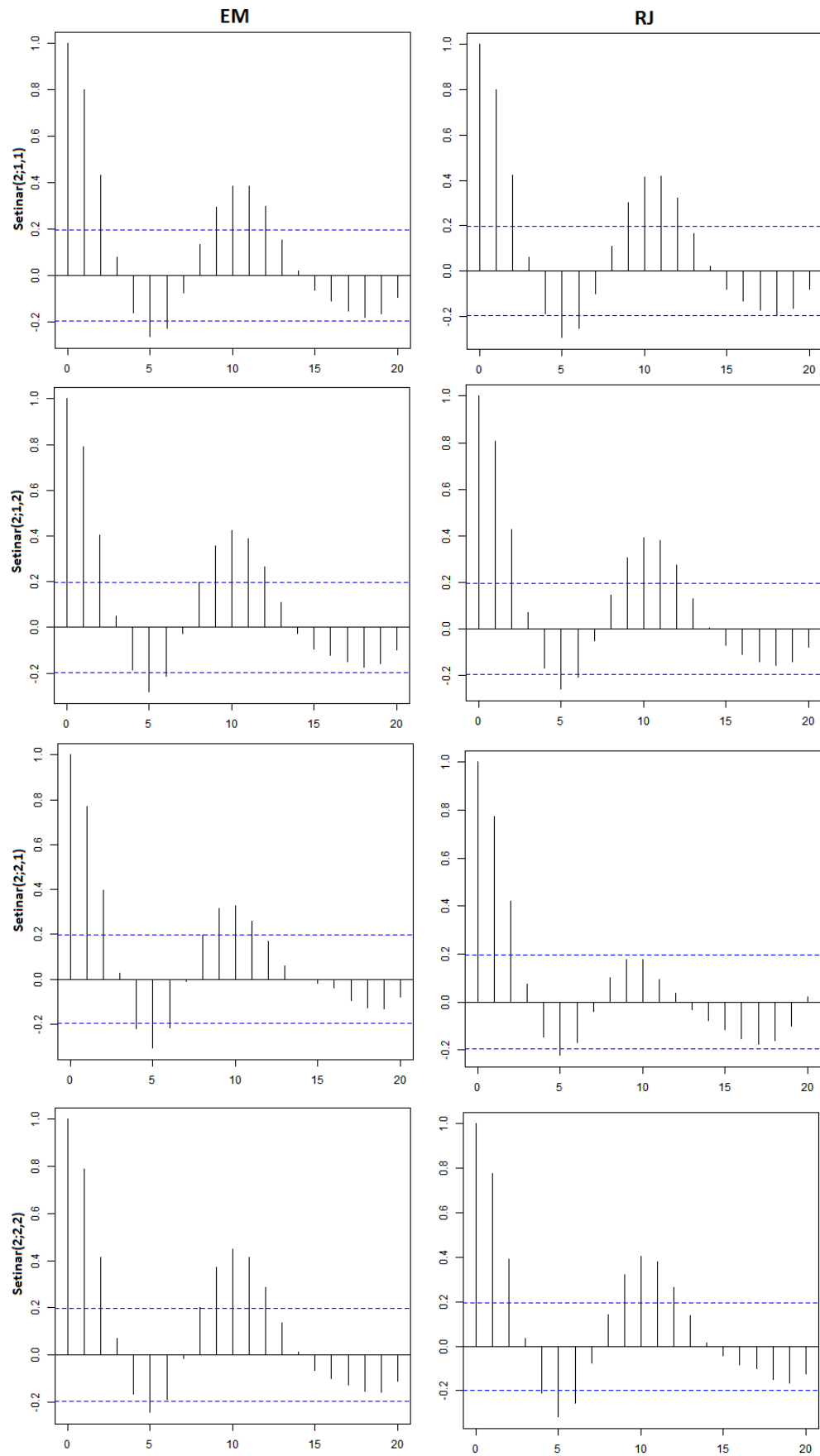


Figura 2.13: FAC da série de Boulder.

Tabela 2.8: Seleção do modelo, usando BIC, DIC e RJMCMC

Modelo	Wölfer		
	BIC	%	DIC
(1,1)	721.05	14	723.00
(1,2)	612.74	36	521.33
(2,1)	501.24	40	486.20
(2,2)	760.12	10	871.02

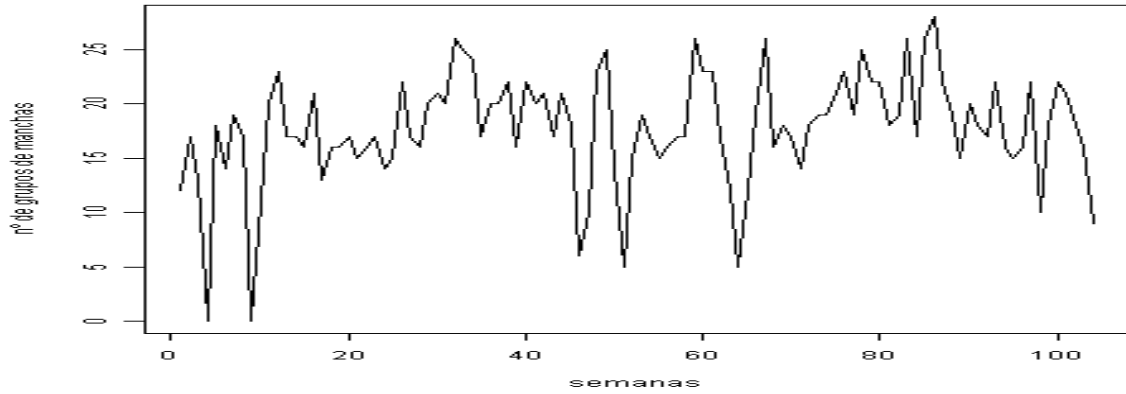
Tabela 2.9: Estimativas de θ para a série de Wölfer.

Modelo	(1,1)		(1,2)		(2,1)		(2,2)	
	EM	MCMC	EM	MCMC	EM	MCMC	EM	MCMC
θ_i								
$\alpha_1^{(1)}$	0.2996	0.2980	0.2815	0.2872	0.2511	0.2812	0.2312	0.2013
$\alpha_2^{(1)}$	-	-	-	-	0.2001	0.1124	0.2200	0.1446
$\alpha_1^{(2)}$	0.3101	0.1526	0.2012	0.2001	-	-	0.3101	0.2813
$\alpha_2^{(2)}$	-	-	0.1811	0.1512	0.1901	0.1201	0.1121	0.1342
$\lambda^{(1)}$	15.3217	15.2110	15.008	15.2364	13.0023	15.3701	13.1001	14.1529
$\lambda^{(2)}$	10.2311	13.4422	10.2012	11.9518	11.0201	12.0081	9.2100	10.1211

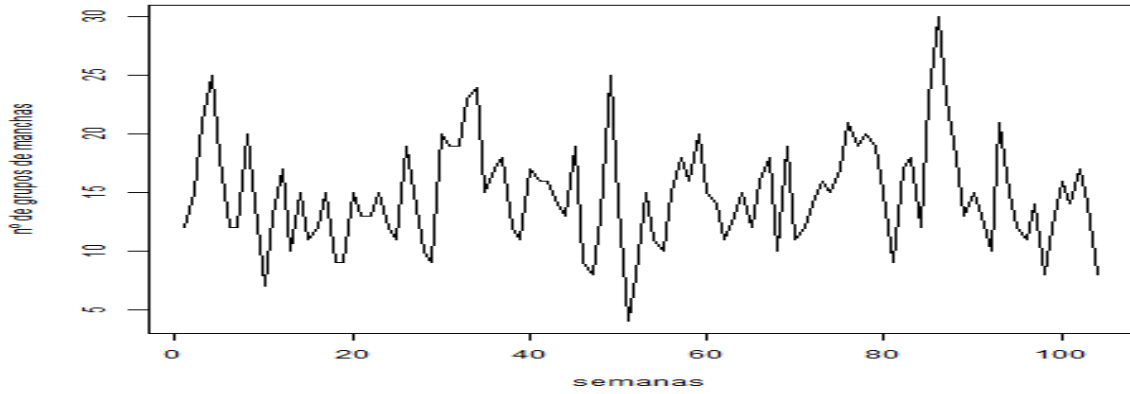
2.5.2 Aplicação 2

Os dados aqui utilizados são disponíveis on-line no site National Geophysical Data Center (<http://www.ngdc.noaa.gov/>) e foram analisados por Silva *et al.* (2005). A série temporal consiste no número total de grupos de manchas solares por semana no período de 2 anos (1990-1991) resultando um total de 104 observações registradas por dois observatórios solares: Centro nacional de dados geofísicos de Boulder (Colorado, EUA) e Palehua Solar Obsevatory (Hawai, EUA). Interpreta-se o número de grupo de manchas solares numa semana como sendo a soma entre o número de grupos de manchas existentes na semana anterior, com probabilidade α , e o número de grupos de manchas que se formaram na semana corrente.

A Figura 2.14 representa as séries temporais de Boulder e Palehua, que podem ser encontradas na íntegra no Apêndice C.1.



(a) Palehua



(b) Boulder

Figura 2.14: Série temporal do número total de grupos de manchas solares no período de 1 de Janeiro de 1990 a 31 de Dezembro de 1991.

Procedendo como anteriormente, foram consideradas todas as combinações dos modelos $\text{SETINAR}(2; p^{(1)}, p^{(2)})$, $1 \leq p^{(k)} \leq 2$, $k = 1, 2$. Para a série Palehua usou-se como limiar $r = 14$ e como desfasamento $d = 2$, para a série temporal do observatório de Boulder admitiu-se limiar $r = 16$ e desfasamento $d = 2$. A escolha do modelo com melhor ajuste, para ambas as séries, pode ser verificada na tabela 2.10 onde são apresentados os valores do BIC, DIC e percentagens da escolha do modelo obtidas pelo RJ. O modelo escolhido para a série de Palehua foi o $\text{SETINAR}(2; 1, 2)$ e para a série de Boulder o modelo $\text{SETINAR}(2; 2, 1)$. Em ambos os casos os 3 métodos

utilizados para a seleção de modelos (DIC,BIC e %RJ) corroboram a escolha.

Nas Tabelas 2.11 e 2.12, respectivamente, estão as médias amostrais das estimativas obtidas para o parâmetro θ para as séries de Palehua e Boulder, através do uso dos algoritmos EM e RJMCMC obtidos nos capítulos anteriores.

Tabela 2.10: Seleção do modelo, usando BIC, DIC e RJMCMC

Modelo	Palehua			Boulder		
	BIC	%	DIC	BIC	%	DIC
(1,1)	755.95	18	823.12	685.52	20	656.20
(1,2)	625.38	32	755.23	421.21	23	645.51
(2,1)	707.53	24	781.20	402.87	47	591.71
(2,2)	703.28	26	808.17	587.33	10	725.27

Tabela 2.11: Média amostral das estimativas de θ para a série Palehua

Modelo	(1,1)		(1,2)		(2,1)		(2,2)	
	EM	RJ	EM	RJ	EM	RJ	EM	RJ
θ_i								
$\alpha_1^{(1)}$	0.5819	0.4253	0.3820	0.3520	0.2765	0.2845	0.2415	0.2911
$\alpha_2^{(1)}$	-	-	-	-	0.1340	0.1624	0.1881	0.1004
$\alpha_1^{(2)}$	0.4486	0.4102	0.3789	0.2213	0.2469	0.3021	0.3210	0.2315
$\alpha_2^{(2)}$	-	-	0.1112	0.1361	-	-	0.1001	0.1131
$\lambda^{(1)}$	7.4922	11.2254	9.2013	10.5833	10.526	9.0020	9.1200	10.1421
$\lambda^{(2)}$	7.7622	3.2145	3.0214	3.5515	3.1420	2.9630	3.0123	3.1516

Tabela 2.12: Média amostral das estimativas de θ para a série Boulder

Modelo	(1,1)		(1,2)		(2,1)		(2,2)	
	EM	RJ	EM	RJ	EM	RJ	EM	RJ
θ_i								
$\alpha_1^{(1)}$	0.3312	0.3520	0.3341	0.3121	0.2362	0.2822	0.2011	0.1108
$\alpha_2^{(1)}$	-	-	-	-	0.2122	0.2051	0.1235	0.1012
$\alpha_1^{(2)}$	0.3724	0.3855	0.1627	0.1227	0.3356	0.3014	0.1787	0.2350
$\alpha_2^{(2)}$	-	-	0.2851	0.3350	-	-	0.2541	0.2142
$\lambda^{(1)}$	7.2513	7.9120	7.2143	7.4231	7.8212	7.0012	5.2314	3.0078
$\lambda^{(2)}$	16.2010	11.0003	13.0430	10.8181	11.0210	10.032	7.5512	7.6527

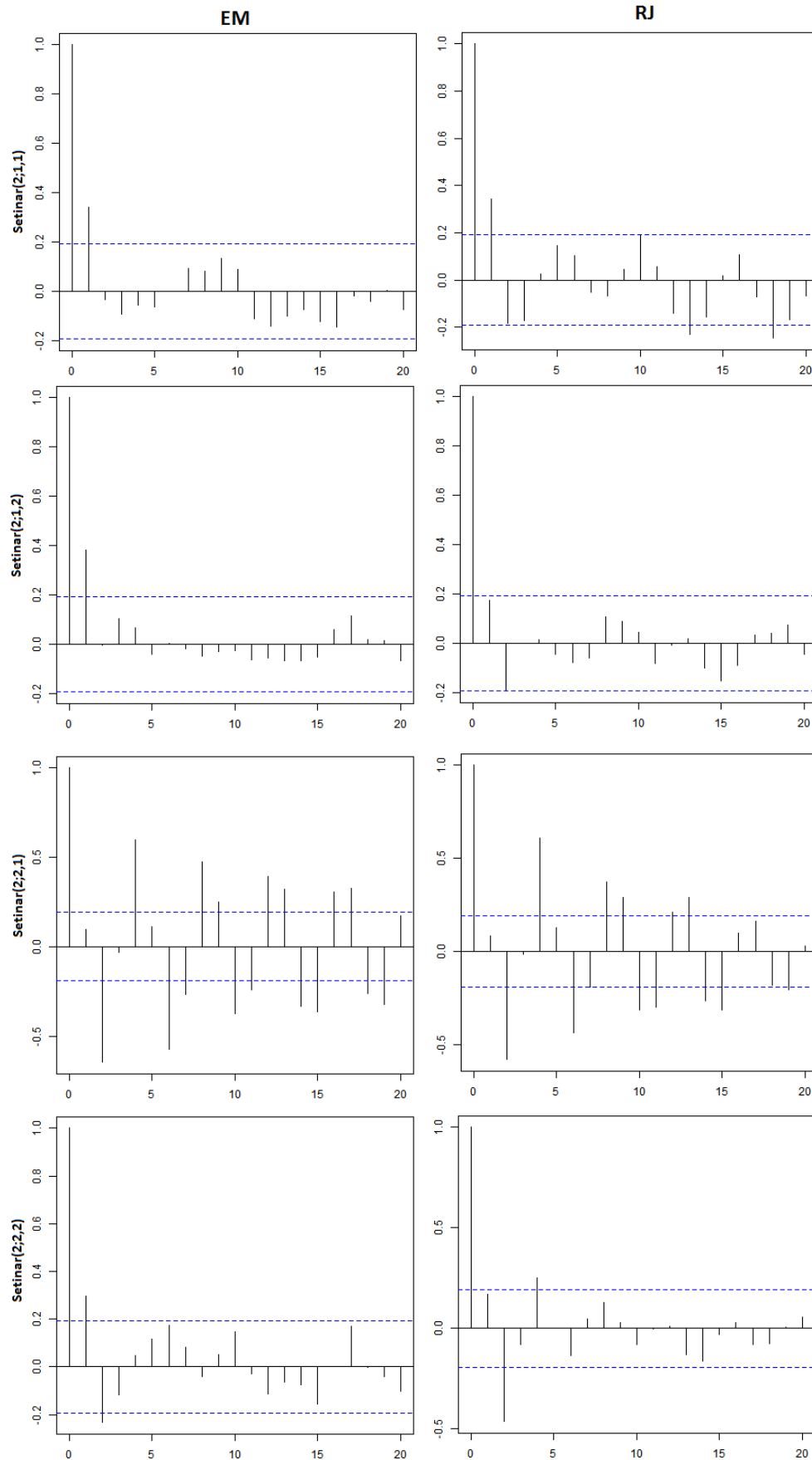


Figura 2.15: FAC da série de Palehua.

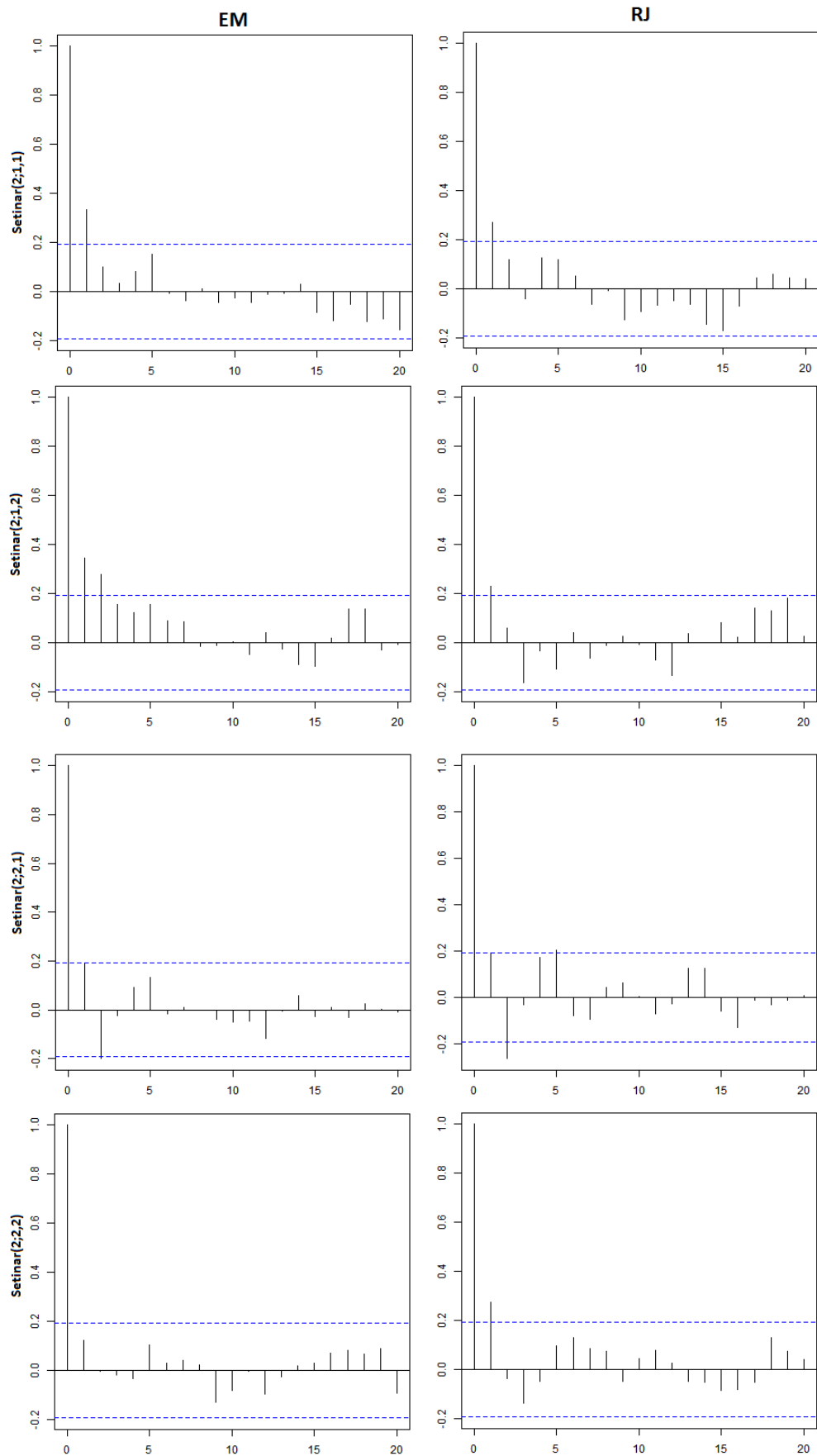


Figura 2.16: FAC da série de Boulder.

2.6 Conclusões

Neste capítulo foi introduzida a classe de modelos autorregressivos de valores inteiros com limiares autoinduzidos designado por $\text{SETINAR}(2; p^{(1)}, p^{(2)})$, a qual apresenta dois regimes cujas componentes autorregressivas de cada ramo são dadas por $p^{(1)}$ e $p^{(2)}$ e as inovações são variáveis aleatórias independentes com distribuição de Poisson. Estabeleceram-se as suas propriedades probabilísticas, média, variância e função de autocovariância do processo; bem como se verificou a existência de um processo estritamente estacionário. Consideraram-se que as inovações eram independentes e identicamente distribuídas, diferentes para cada um dos troços.

Determinaram-se os estimadores de mínimos quadrados condicionais e as suas propriedades assintóticas e as estimativas de máxima verosimilhança foram obtidas através do algoritmo EM. Para completar, fez-se um estudo de simulação onde se analisaram os desempenhos das estimativas obtidas pela metodologia clássica e pela metodologia bayesiana, usando métodos MCMC. No estudo de simulação efetuado, consideraram-se conhecidas as ordens dos modelos de cada regime e também se admitiram desconhecidas, procedendo-se adicionalmente à seleção de modelos, usando o RJMCMC e os critérios BIC e DIC.

Verificou-se a superioridade do método RJMCMC na estimação dos parâmetros e na determinação das ordens de cada um dos regimes do modelo.

Capítulo 3

Processos Autorregressivos de Valores Inteiros Definidos por Limiares Periódicos Autoinduzidos

*“Nothing occurs at random, but everything
for a reason and by necessity.”*

Leucippus - 5th century B.C.

A classe de processos de valores contínuos periodicamente correlacionados foram estudados por Bennett (1958) e Gladyshev (1961) sendo a classe de modelos autorregressivos com estrutura periódica inicialmente estudada por Jones & Brelsford (1967). Monteiro *et al.* (2010) introduziu uma classe de modelos autorregressivos de valores inteiros não negativos com estrutura periódica de primeira ordem e período T - $\text{PINAR}(1)_T$. Monteiro *et al.* (2010) demonstrou a existência e unicidade de um processo estacionário, apresentou expressões para média, variância e função de autocorrelação para a distribuição ciclo-estacionária. A estimação dos parâmetros também foi abordada no trabalho de Monteiro *et al.* (2010). No intuito de explorar os modelos autorregressivos inteiros com estrutura periódica e os modelos **SETINAR** apresentados no capítulo anterior, desenvolveu-se um modelo autorregressivo de valores inteiros com

limiares periódicos autoinduzidos (PSETINAR¹) o qual introduzir-se-á neste capítulo. Na Secção 3.1 apresenta-se o modelo, a existência de uma distribuição estacionária e a expressão para a média. A estimação dos parâmetros é abordada na Secção 3.2 e um estudo simulado é apresentado na Secção 3.3. Um estudo de predição de novos valores é abordado na Secção 3.4 e conclui-se o capítulo com aplicação a dados reais na Secção 3.5.

3.1 O Modelo PSETINAR(2; 1, 1)_T

A classe dos modelos autorregressivos de valores inteiros de estrutura periódica com limiares autoinduzidos de primeira ordem em cada ramo e período T , designado por PSETINAR(2; 1, 1)_T, é definida por

$$X_t = \begin{cases} \alpha_j^{(1)} \circ X_{t-1} + Z_t^{(1)}, & \text{se } X_{t-d} \leq R_t \\ \alpha_j^{(2)} \circ X_{t-1} + Z_t^{(2)}, & \text{se } X_{t-d} > R_t \end{cases}, \quad t \in \mathbb{N}_0, \quad (3.1)$$

com $R_t = r_j$, $t = j + sT$, $j = 1, \dots, T$ e $s \in \mathbb{N}_0$. Para o período j tem-se que

$$X_{j+sT} = (\alpha_j^{(1)} \circ X_{j+sT-1} + Z_{j+sT}^{(1)})I_j^{(1)} + (\alpha_j^{(2)} \circ X_{j+sT-1} + Z_{j+sT}^{(2)})I_j^{(2)},$$

onde $I_j^{(k)}$, $k = \{1, 2\}$, é a função indicatriz definida por

$$I_j^{(1)} := \begin{cases} 1, & X_{j+sT-d} \leq r_j \\ 0, & X_{j+sT-d} > r_j \end{cases} \quad e \quad I_j^{(2)} = 1 - I_j^{(1)}.$$

A equação (3.1) admite a representação

$$X_t = \phi_t \circ X_{t-1} + Z_t,$$

¹do inglês - *Periodic Self-Exciting Threshold INteger-valued Autoregressive*.

sendo $Z_t = Z_j^{(1)}I_j^{(1)} + Z_j^{(2)}I_j^{(2)}$, $\phi_t = \alpha_j \equiv \alpha_j^{(1)}I_j^{(1)} + \alpha_j^{(2)}I_j^{(2)}$ tal que $\alpha_j \in (0, 1)$, $t = j + sT$, $j = 1, \dots, T$, $s \in \mathbb{N}$. O operador *thinning* binomial “ \circ ” é definido por

$$\phi_t \circ X_{t-1} \stackrel{d}{=} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_j^{(1)})I_j^{(1)} + \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_j^{(2)})I_j^{(2)},$$

sendo $(U_{i,t}(\alpha_j^{(1)}))$ e $(U_{i,t}(\alpha_j^{(2)}))$, $i = 1, 2, \dots$, sucessões periódicas de variáveis aleatórias de Bernoulli com probabilidade de sucesso $P(U_{i,t}(\alpha_j^{(1)}) = 1) = \alpha_j^{(1)}$ e $P(U_{i,t}(\alpha_j^{(2)}) = 1) = \alpha_j^{(2)}$. O processo das inovações (Z_t) é uma sucessão periódica de variáveis aleatórias independentes com distribuição de Poisson de média v_t , $Z_t \sim Po(v_t)$, com $v_t = \lambda_j$ para $t = j + sT$, $j = 1, \dots, T$, $s \in \mathbb{N}_0$, sendo para cada t , independentes de X_{t-1} e de $\alpha_t \circ X_{t-1}$. Após alguns cálculos pode-se escrever o modelo PSETINAR(2; 1, 1)_T na forma de um processo INAR(1) multivariado.

$$\left\{ \begin{array}{l} X_{1+sT} = \alpha_1 \circ X_{1+sT-1} + Z_{1+sT} \\ X_{2+sT} = \alpha_2 \circ X_{2+sT-1} + Z_{2+sT} \\ \quad = \alpha_2 \circ (\alpha_1 \circ X_{1+sT-1} + Z_{1+sT}) + Z_{2+sT} \\ \quad = (\alpha_2 \alpha_1) \circ X_{1+sT-1} + \alpha_2 \circ Z_{1+sT} + Z_{2+sT} \\ X_{3+sT} = \alpha_3 \circ X_{3+sT-1} + Z_{3+sT} \\ \quad = (\alpha_3 \alpha_2 \alpha_1) \circ X_{1+sT-1} + \alpha_3 \alpha_2 \circ Z_{1+sT} + \alpha_3 \circ Z_{2+sT} + Z_{3+sT} \\ \vdots \\ X_{T+sT} = (\alpha_T \alpha_{T-1} \cdots \alpha_2 \alpha_1) \circ X_{1+sT-1} + \{(\alpha_T \alpha_{T-1} \cdots \alpha_2) \circ Z_{1+sT} \\ \quad + (\alpha_T \alpha_{T-1} \cdots \alpha_3) \circ Z_{2+sT} + \cdots + \alpha_T \circ Z_{(T-1)+sT} + Z_{T+sT} \\ \quad = \left(\prod_{l=0}^{T-1} \alpha_{T-l} \right) \circ X_{1+sT-1} + \sum_{i=1}^{T-1} \left(\prod_{l=0}^{T-1} \alpha_{i-l} \right) \circ Z_{T-i+sT} + Z_{T+sT} \end{array} \right. .$$

De uma forma geral,

$$X_{j+sT} = \left(\prod_{l=0}^{j-1} \alpha_{j-l} \right) \circ X_{1+sT-1} + \sum_{i=1}^{j-1} \left(\prod_{l=0}^{i-1} \alpha_{j-l-i} \right) \circ Z_{i+sT} + Z_{j+sT},$$

para $j = 1, \dots, T$ e $s \in \mathbb{N}_0$. A equação (3.1) admite a seguinte representação vetorial

$$\mathbf{Y}_t = \mathbf{A} \circ \mathbf{Y}_{t-1} + \zeta_t, \quad (3.2)$$

com $\mathbf{Y}_t = [X_{1+tT} \ X_{2+tT} \ \dots \ X_{T+tT}]'$ e $\zeta_t = \mathbf{B} \circ \mathbf{Z}_t$,

$$\mathbf{A}_{T \times T} = \begin{pmatrix} 0 & 0 & \dots & \alpha_1 \\ 0 & 0 & \dots & \alpha_1 \alpha_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \prod_{j=0}^{T-1} \alpha_{T-j} \end{pmatrix}$$

e

$$\mathbf{B}_{T \times T} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \alpha_2 & 1 & \dots & 0 & 0 \\ \alpha_3 \alpha_2 & \alpha_3 & \dots & 0 & 0 \\ \vdots & \ddots & \dots & \dots & \vdots \\ \prod_{i=0}^{T-1-2} \alpha_{T-1-i} & \prod_{i=0}^{T-1-3} \alpha_{T-1-i} & \dots & 1 & 0 \\ \prod_{i=0}^{T-1-1} \alpha_{T-i} & \prod_{i=0}^{T-1-2} \alpha_{T-i} & \dots & \alpha_T & 1 \end{pmatrix}.$$

Nos resultados que se seguem, introduz-se a distribuição (ciclo)estacionária de (X_t) para cada um dos T períodos.

Proposição 3.1. *Seja $0 < \alpha_t < 1$, $j = 1, \dots, T$ e $0 < P(\mathbf{Z}_t = \mathbf{0}) < 1$. Então qualquer solução (\mathbf{Y}_t) de (3.2) é uma cadeia de Markov irredutível e aperiódica.*

Demonstração. A partir da equação (3.2) verifica-se que \mathbf{Y}_t é uma cadeia de Markov irredutível e aperiódica. Para mostrar a irredutibilidade equivale a verificar que todos

os estados da cadeia comunicam-se entre si, tem-se

$$\begin{aligned}
 P(\mathbf{m}, \mathbf{n}) &= P(\mathbf{A} \circ \mathbf{Y}_{t-1} + \zeta_t = \mathbf{n} | \mathbf{Y}_{t-1} = \mathbf{m}) \\
 &= \sum_{i_1=0}^{n_1} \dots \sum_{i_T=0}^{n_T} \prod_{j=1}^T P(Z_{j+tT} = i_j) \times P\left(\prod_{k=1}^j \alpha_k \circ m_T + \sum_{k=1}^{j-1} \left(\prod_{l=0}^{j-k-1} \alpha_{T-l}\right) \circ Z_{j+tT} = \right. \\
 &\quad \left. n_j - i_j \middle| Z_{1+tT} = i_1, \dots, Z_{T+tT} = i_T\right) \\
 &\geq \prod_{j=1}^T P(Z_{j+tT} = n_j) \times P\left(\prod_{k=1}^j \alpha_k \circ m_T + \sum_{k=1}^{j-1} \left(\prod_{l=0}^{j-k-1} \alpha_{T-l}\right) \circ Z_{j+tT} = \right. \\
 &\quad \left. n_j - i_j \middle| Z_{1+tT} = i_1, \dots, Z_{T+tT} = i_T\right) \\
 &\geq \prod_{j=1}^T \prod_{k=1}^{j-1} P(Z_{j+tT} = n_j) \times \left(1 - \prod_{k=1}^j \alpha_j\right)^{m_T} \left(1 - \prod_{k=1}^j \alpha_k\right)^{n_1} \times \dots \times \left(1 - \prod_{k=T}^j \alpha_k\right)^{n_{T-1}} \\
 &> 0,
 \end{aligned}$$

com $\prod_{k=i}^j \alpha_k = 0$, $i > j$. Assim,

$$\begin{aligned}
 P(\mathbf{m}, \mathbf{0}) &= \prod_{j=1}^T P(Z_{j+tT} = 0) \left(1 - \prod_{k=1}^j \alpha_k\right)^{m_T} \\
 P(\mathbf{0}, \mathbf{n}) &= \prod_{j=1}^T P(Z_{j+tT} = n_j) \left(1 - \prod_{l=0}^{j-1} \alpha_{j-l}\right)^{n_{j-l}} > 0,
 \end{aligned}$$

o que implica que \mathbf{Y}_t é irredutível. Uma vez que

$$P(\mathbf{0}, \mathbf{0}) = \prod_{j=1}^T P(Z_{j+sT} = 0) > 0,$$

então, \mathbf{Y}_t é uma cadeia de Markov aperiódica (Franke & Subba Rao, 1995). ■

A existência quase certa de um único processo estritamente estacionário para o modelo $\text{PSETINAR}(2; 1, 1)_T$ segue a proposta de Latour (1997) para o processo $\text{MGINAR}(1)$. Para tal considera-se as condições da Proposição 3.1 em Latour (1997).

- \mathbf{A} é um operador tal que $\det(\mathbf{I}z - \mathbf{A}) = 0$ tem todas as raízes do polinómio

característico dentro do círculo unitário, ou seja o maior valor próprio η de \mathbf{A} é menor que 1;

- ζ_t é uma sequência i.i.d. de vetores aleatórios de média finita $E[\zeta_t] = \mathbf{B}[\lambda_1, \dots, \lambda_T]'$.

Proposição 3.2. *O processo (\mathbf{Y}_t) em (3.2) é uma cadeia de Markov irredutível e aperiódica. Se $E[\zeta_t] < \infty$ e o maior valor próprio η de \mathbf{A} é menor que 1, então existe um processo estritamente estacionário que satisfaz (3.2).*

Demonstração. Pela Proposição 3.1 tem-se que a cadeia de Markov é irredutível e aperiódica. O vetor ζ_t é independente de ζ_{t+j} por serem funções dos vetores \mathbf{Z}_t e \mathbf{Z}_{t+j} que são independentes entre si. A variância é dada por $V[\zeta_t] = \mathbf{B} \sum_Z \mathbf{B}'$ onde $\sum_Z := \text{Cov}(\mathbf{Z}_t, \mathbf{Z}_{t+j}) = \text{diag}[\lambda_1, \dots, \lambda_T]'$. Resta verificar que ζ_t é uma sucessão i.i.d. de vetores aleatórios e independentes de \mathbf{Y}_t . Note-se que

$$\begin{aligned}
 P(\zeta_t = \mathbf{k}) &= P(\mathbf{B} \circ Z_t = \mathbf{k}) \\
 &= P \left(\begin{array}{l} Z_{1+tT} = k_1, \alpha_2 \circ Z_{1+tT} + Z_{2+tT} = k_2, \alpha_3 \alpha_2 \circ Z_{1+tT} + \alpha_3 \circ Z_{2+tT} + \\ Z_{3+tT} = k_3, \dots, \sum_{i=1}^{T-1} \left(\prod_{m=0}^{T-1-i} \alpha_{s-m} \right) \circ Z_{i+tT} + Z_{T+tT} = k_T \end{array} \right) \\
 &= P \left(\begin{array}{l} Z_1 = k_1, \alpha_2 \circ Z_1 + Z_2 = k_2, \alpha_3 \alpha_2 \circ Z_1 + \alpha_3 \circ Z_2 + Z_3 = k_3, \\ \dots, \sum_{i=1}^{T-1} \left(\prod_{m=0}^{T-1-i} \alpha_{t-m} \right) \circ Z_i + Z_T = k_T \end{array} \right) \\
 &= P \left(\begin{array}{l} Z_{1+hT} = k_1, \alpha_2 \circ Z_{1+hT} + Z_{2+hT} = k_2, \alpha_3 \alpha_2 \circ Z_{1+hT} + \alpha_3 \circ Z_{2+hT} + \\ Z_{3+hT} = k_3, \dots, \sum_{i=1}^{T-1} \left(\prod_{m=0}^{T-1-i} \alpha_{T-m} \right) \circ Z_{i+hT} + Z_{T+hT} = k_T \end{array} \right) \\
 &= P(\mathbf{B} \circ Z_h = \mathbf{k}) \\
 &= P(\zeta_h = \mathbf{k}),
 \end{aligned}$$

logo $P(\zeta_t = \mathbf{k}) = P(\mathbf{B} \circ Z_t = \mathbf{k}) \forall t, h$, então ζ_t é i.i.d. Desta forma sendo (ζ_t) uma sucessão i.i.d de vetores aleatórios e independentes de \mathbf{Y}_t , $E[\zeta_t] < \infty$ e sendo o maior valor próprio da matriz \mathbf{A} menor que 1 pode-se verificar que existe um processo estritamente estacionário que satisfaz (3.2). ■

Lema 3.1. Para o modelo $PSETINAR(2; 1, 1)_T$ com $t = j + sT$, $j = 1, \dots, T$ e $s \in \mathbf{N}_0$ a média é dada pela expressão

$$\begin{aligned} E[X_t] &= \sum_{k_1=1}^2 \sum_{k_2=1}^2 \dots \sum_{k_j=1}^2 \left[\alpha_1^{(k_1)} \alpha_2^{(k_2)} \dots \alpha_j^{(k_j)} \times u_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} \times p_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} \right] \\ &+ \sum_{l=1}^j \lambda_l \left(\sum_{k_{l+1}=1}^2 \dots \sum_{k_j=1}^2 \alpha_{l+1}^{(k_{l+1})} \dots \alpha_j^{(k_j)} \times p_{l+1,\dots,j}^{k_{l+1},\dots,k_j} \right), \end{aligned}$$

onde

$$u_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} := E \left[X_{sT} | X_{1+sT-d} \in r_1^{(k_1)}, X_{2+sT-d} \in r_2^{(k_2)}, \dots, X_{j+sT-d} \in r_j^{(k_j)} \right],$$

$$p_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} := P \left[X_{1+sT-d} \in r_1^{(k_1)}, X_{2+sT-d} \in r_2^{(k_2)}, \dots, X_{j+sT-d} \in r_j^{(k_j)} \right],$$

para $k_1, k_2, \dots, k_j = 1, 2$ e onde r_j representa o ramo do período j , isto é,

$$r_j = \begin{cases} \text{ramo } r_j^{(1)} & \text{se } X_{j+sT-d} \leq r_j \\ \text{ramo } r_j^{(2)} & \text{se } X_{j+sT-d} > r_j \end{cases}.$$

Demonstração. Note-se que

$$E[X_{j+sT}] = \underbrace{E \left[\left(\prod_{l=0}^{j-1} \alpha_{j-l} \right) \circ X_{1+sT-1} \right]}_I + \underbrace{E \left[\sum_{i=1}^{j-1} \left(\prod_{l=0}^{j-i-1} \alpha_{j-l} \right) \circ Z_{i+sT} \right]}_{II} + \underbrace{E[Z_{j+sT}]}_{III}. \quad (3.3)$$

O termo I de (3.3) é dado por

$$\begin{aligned} E \left[\left(\prod_{l=0}^{j-1} \alpha_{j-l} \right) \circ X_{1+sT-1} \right] &= E \left[(\alpha_1 \dots \alpha_j) \circ X_{sT} \right] \\ &= E \left[\left(\alpha_1^{(1)} I_1^{(1)} + \alpha_1^{(2)} I_1^{(2)} \right) \times \left(\alpha_2^{(1)} I_2^{(1)} + \alpha_2^{(2)} I_2^{(2)} \right) \times \dots \times \left(\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)} \right) \circ X_{sT} \right] \end{aligned}$$

$$\begin{aligned}
&= E \left[\left(\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(1)} I_2^{(1)} + \alpha_1^{(1)} I_1^{(1)} \alpha_2^{(2)} I_2^{(2)} + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(1)} I_2^{(1)} + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(2)} I_2^{(2)} \right) \times \cdots \times \right. \right. \\
&\quad \left. \left. \times \left(\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)} \right) \right) \circ X_{sT} \right] \\
&\quad \vdots \\
&= E \left[\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(1)} I_j^{(1)} + \alpha_1^{(1)} I_1^{(1)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(1)} I_j^{(1)} + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(1)} I_2^{(1)} \cdots \right. \right. \\
&\quad \alpha_j^{(1)} I_j^{(1)} + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(1)} I_j^{(1)} + \cdots + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(2)} I_j^{(2)} + \alpha_1^{(1)} I_1^{(1)} \\
&\quad \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(2)} I_j^{(2)} + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(2)} I_j^{(2)} + \alpha_1^{(2)} I_1^{(2)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(2)} I_j^{(2)} \Big) \circ X_{sT} \Big] \\
&= E \left[\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(1)} I_j^{(1)} \right) \circ X_{sT} \right] + E \left[\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(1)} I_j^{(1)} \right) \circ X_{sT} \right] \\
&\quad + E \left[\left(\alpha_1^{(2)} I_1^{(2)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(1)} I_j^{(1)} \right) \circ X_{sT} \right] + E \left[\left(\alpha_1^{(2)} I_1^{(2)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(1)} I_j^{(1)} \right) \circ X_{sT} \right] \\
&\quad + \cdots + E \left[\left(\alpha_1^{(2)} I_1^{(1)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(2)} I_j^{(2)} \right) \circ X_{sT} \right] + \\
&\quad + E \left[\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(2)} I_j^{(2)} \right) \circ X_{sT} \right] + E \left[\left(\alpha_1^{(2)} I_1^{(2)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(2)} I_j^{(2)} \right) \circ X_{sT} \right] \\
&\quad + E \left[\left(\alpha_1^{(2)} I_1^{(2)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(2)} I_j^{(2)} \right) \circ X_{sT} \right].
\end{aligned}$$

A cada uma das 2^T parcelas anteriores, aplicam-se as propriedades do operador *thinning* e do valor esperado tal como

$$\begin{aligned}
&E \left[\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(1)} I_2^{(1)} \cdots \alpha_j^{(1)} I_j^{(1)} \right) \circ X_{sT} \right] = \alpha_1^{(1)} \alpha_2^{(1)} \cdots \alpha_j^{(1)} E \left[X_{sT} \middle| X_{1+sT-d} \leq r_1, \right. \\
&\quad \left. X_{2+sT-d} \leq r_2, \dots, X_{j+sT-d} \leq r_j \right] P \left[X_{1+sT-d} \leq r_1, X_{2+sT-d} \leq r_2, \dots, X_{j+sT-d} \leq r_j \right] \\
&= \alpha_1^{(1)} \alpha_2^{(1)} \cdots \alpha_j^{(1)} u_{1,2,\dots,j}^{(1,1,\dots,1)} p_{1,2,\dots,j}^{(1,1,\dots,1)} \\
&E \left[\left(\alpha_1^{(1)} I_1^{(1)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(1)} I_j^{(1)} \right) \circ X_{sT} \right] = \alpha_1^{(1)} \alpha_2^{(2)} \cdots \alpha_j^{(1)} E \left[X_{sT} \middle| X_{1+sT-d} \leq r_1, \right. \\
&\quad \left. X_{2+sT-d} > r_2, \dots, X_{j+sT-d} \leq r_j \right] P \left[X_{1+sT-d} \leq r_1, X_{2+sT-d} > r_2, \dots, X_{j+sT-d} \leq r_j \right] \\
&= \alpha_1^{(1)} \alpha_2^{(2)} \cdots \alpha_j^{(1)} u_{1,2,\dots,j}^{(1,2,\dots,1)} p_{1,2,\dots,j}^{(1,2,\dots,1)} \\
&\quad \vdots \\
&E \left[\left(\alpha_1^{(2)} I_1^{(2)} \alpha_2^{(2)} I_2^{(2)} \cdots \alpha_j^{(2)} I_j^{(2)} \right) \circ X_{sT} \right] = \alpha_1^{(2)} \alpha_2^{(2)} \cdots \alpha_j^{(2)} E \left[X_{sT} \middle| X_{1+sT-d} > r_1, \right. \\
&\quad \left. X_{2+sT-d} > r_2, \dots, X_{j+sT-d} > r_j \right] P \left[X_{1+sT-d} > r_1, X_{2+sT-d} > r_2, \dots, X_{j+sT-d} > r_j \right] \\
&= \alpha_1^{(2)} \alpha_2^{(2)} \cdots \alpha_j^{(2)} u_{1,2,\dots,j}^{(2,2,\dots,2)} p_{1,2,\dots,j}^{(2,2,\dots,2)}.
\end{aligned}$$

Logo, tem-se

$$\begin{aligned}
 E \left[\left(\prod_{l=0}^{j-1} \alpha_{j-l} \right) \circ X_{1+sT-1} \right] &= \\
 &= \left(\alpha_1^{(1)} \alpha_2^{(1)} \cdots \alpha_j^{(1)} \right) u_{1,2,\dots,j}^{(1,1,\dots,1)} p_{1,2,\dots,j}^{(1,1,\dots,1)} + \left(\alpha_1^{(1)} \alpha_2^{(2)} \cdots \alpha_j^{(1)} \right) u_{1,2,\dots,j}^{(1,2,\dots,1)} p_{1,2,\dots,j}^{(1,2,\dots,1)} \\
 &+ \left(\alpha_1^{(1)} \alpha_2^{(2)} \cdots \alpha_j^{(1)} \right) u_{1,2,\dots,j}^{(1,2,\dots,1)} p_{1,2,\dots,j}^{(1,2,\dots,1)} + \left(\alpha_1^{(1)} \alpha_2^{(2)} \cdots \alpha_j^{(2)} \right) u_{1,2,\dots,j}^{(1,2,\dots,2)} p_{1,2,\dots,j}^{(1,2,\dots,2)} \\
 &+ \cdots + \left(\alpha_1^{(2)} \alpha_2^{(1)} \cdots \alpha_j^{(1)} \right) u_{1,2,\dots,j}^{(2,1,\dots,1)} p_{1,2,\dots,j}^{(2,1,\dots,1)} + \left(\alpha_1^{(2)} \alpha_2^{(1)} \cdots \alpha_j^{(2)} \right) u_{1,2,\dots,j}^{(2,1,\dots,2)} p_{1,2,\dots,j}^{(2,1,\dots,2)} \\
 &+ \left(\alpha_1^{(2)} \alpha_2^{(2)} \cdots \alpha_j^{(1)} \right) u_{1,2,\dots,j}^{(2,2,\dots,1)} p_{1,2,\dots,j}^{(2,2,\dots,1)} + \left(\alpha_1^{(2)} \alpha_2^{(2)} \cdots \alpha_j^{(2)} \right) u_{1,2,\dots,j}^{(2,2,\dots,2)} p_{1,2,\dots,j}^{(2,2,\dots,2)} \\
 &= \sum_{k_1=1}^2 \cdots \sum_{k_j=1}^2 \alpha_1^{(k_1)} \alpha_2^{(k_2)} \cdots \alpha_j^{(k_j)} u_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} p_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)}.
 \end{aligned}$$

O termo II de (3.3) é dado por

$$\begin{aligned}
 E \left[\sum_{i=1}^{j-1} \left(\prod_{l=0}^{j-i-1} \alpha_{j-l} \right) \circ Z_{i+sT} \right] &= E \left[(\alpha_2 \alpha_3 \cdots \alpha_j) \circ Z_{1+sT} + (\alpha_3 \cdots \alpha_j) \circ Z_{2+sT} + \cdots \right. \\
 &\quad \left. + (\alpha_{j-1} \alpha_j) \circ Z_{(j-2)+sT} + (\alpha_j) \circ Z_{(j-1)+sT} \right] \\
 &= E \left[\left((\alpha_2^{(1)} I_2^{(1)} + \alpha_2^{(2)} I_2^{(2)}) \times (\alpha_3^{(1)} I_3^{(1)} + \alpha_3^{(2)} I_3^{(2)}) \times \cdots \times (\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}) \right) \right. \\
 &\quad \left. \circ Z_{1+sT} \right] + E \left[\left((\alpha_3^{(1)} I_3^{(1)} + \alpha_3^{(2)} I_3^{(2)}) \times \cdots \times (\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}) \right) \circ Z_{1+sT} \right] + \\
 &+ \cdots + E \left[\left((\alpha_{j-1}^{(1)} I_{j-1}^{(1)} + \alpha_{j-1}^{(2)} I_{j-1}^{(2)}) \times (\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}) \right) \circ Z_{(j-2)+sT} \right] \\
 &+ E \left[(\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}) \circ Z_{(j-1)+sT} \right].
 \end{aligned}$$

Após alguma álgebra simples tem-se

$$\begin{aligned}
 E \left[\sum_{i=1}^{j-1} \left(\prod_{l=0}^{j-i-1} \alpha_{j-l} \right) \circ Z_{i+sT} \right] &= \lambda_1 \left\{ \sum_{k_2=1}^2 \cdots \sum_{k_j=1}^2 \left[\alpha_2^{(k_2)} \alpha_3^{(k_3)} \cdots \alpha_j^{(k_j)} \times p_{2,3,\dots,j}^{(k_2,k_3,\dots,k_j)} \right] \right\} \\
 &+ \lambda_2 \left\{ \sum_{k_3=1}^2 \cdots \sum_{k_j=1}^2 \left[\alpha_3^{(k_3)} \cdots \alpha_j^{(k_j)} \times p_{2,3,\dots,j}^{(k_2,k_3,\dots,k_j)} \right] \right\} + \cdots + \lambda_{j-2} \left\{ \sum_{k_{j-1}=1}^2 \sum_{k_j=1}^2 \left[\alpha_{j-1}^{(k_{j-1})} \right. \right. \\
 &\quad \left. \left. \alpha_j^{(k_j)} \times p_{j-1,j}^{(k_{j-1},k_j)} \right] \right\} + \lambda_{j-1} \sum_{k_j=1}^2 \left[\alpha_j^{(k_j)} \times p_j^{(k_j)} \right] \\
 E \left[\sum_{i=1}^{j-1} \left(\prod_{l=0}^{j-i-1} \alpha_{j-l} \right) \circ Z_{i+sT} \right] &= \sum_{l=1}^{j-1} \lambda_l \left\{ \sum_{k_{l+1}=1}^2 \cdots \sum_{k_j=1}^2 \alpha_{l+1}^{(k_{l+1})} \cdots \alpha_j^{(k_j)} \times p_{l+1,\dots,j}^{(k_{l+1},\dots,k_j)} \right\}.
 \end{aligned}$$

O termo *III* de (3.3) é dado por

$$E[Z_{j+sT}] = \lambda_j.$$

Assim $E[X_t] = E[X_{j+sT}] = I + II + III$ é dada pela expressão

$$\begin{aligned} E[X_t] &= \sum_{k_1=1}^2 \sum_{k_2=1}^2 \dots \sum_{k_j=1}^2 \left[\alpha_1^{(k_1)} \alpha_2^{(k_2)} \dots \alpha_j^{(k_j)} \times u_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} \times p_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} \right] + \\ &\quad + \sum_{l=1}^{j-1} \lambda_l \left(\sum_{k_{l+1}=1}^2 \dots \sum_{k_j=1}^2 \alpha_{l+1}^{(k_{l+1})} \dots \alpha_j^{(k_j)} \times p_{l+1,\dots,j}^{(k_{l+1},\dots,k_j)} \right) + \lambda_j \\ &= \sum_{k_1=1}^2 \sum_{k_2=1}^2 \dots \sum_{k_j=1}^2 \left[\alpha_1^{(k_1)} \alpha_2^{(k_2)} \dots \alpha_j^{(k_j)} \times u_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} \times p_{1,2,\dots,j}^{(k_1,k_2,\dots,k_j)} \right] \\ &\quad + \sum_{l=1}^j \lambda_l \left(\sum_{k_{l+1}=1}^2 \dots \sum_{k_j=1}^2 \alpha_{l+1}^{(k_{l+1})} \dots \alpha_j^{(k_j)} \times p_{l+1,\dots,j}^{(k_{l+1},\dots,k_j)} \right). \end{aligned}$$

■

3.2 Métodos de Estimação

O objetivo desta secção é aplicar o método dos mínimos quadrados condicionais proposto por Klimko & Nelson (1978) e o método de máxima verossimilhança condicional aos modelos de valores inteiros de limiares autoinduzidos com estrutura periódica ($\text{PSETINAR}(2; 1, 1)_T$) e apresentar as suas propriedades assintóticas. Seja (X_1, \dots, X_n) uma sucessão de variáveis aleatórias que satisfazem (3.1) e $\boldsymbol{\theta} := (\alpha_1^{(1)}, \alpha_1^{(2)}, \lambda_1, \dots, \alpha_T^{(1)}, \alpha_T^{(2)}, \lambda_T)$, o vetor dos parâmetros desconhecidos. Ao longo desta secção assume-se que os limiares R são conhecidos.

3.2.1 Estimadores de Mínimos Quadrados Condicionais

O estimador de MQC é o vetor

$$\hat{\boldsymbol{\theta}}_{MQC} := (\hat{\alpha}_{1,MQC}^{(1)}, \hat{\alpha}_{1,MQC}^{(2)}, \hat{\lambda}_{1,MQC}, \dots, \hat{\alpha}_{T,MQC}^{(1)}, \hat{\alpha}_{T,MQC}^{(2)}, \hat{\lambda}_{T,MQC}),$$

o qual minimiza a expressão

$$Q(\boldsymbol{\theta}) := \sum_{s=0}^{N-1} \sum_{j=1}^T \left(X_{j+sT} - g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \right)^2,$$

onde N é o número de ciclos completos, T é o número de períodos, $\boldsymbol{\theta}_j := (\alpha_j^{(1)}, \alpha_j^{(2)}, \lambda_j)$,

$$g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) = \alpha_j^{(1)} X_{j+sT-1} I_j^{(1)} + \alpha_j^{(2)} X_{j+sT-1} I_j^{(2)} + \lambda_j,$$

e $\boldsymbol{\theta} =: (\boldsymbol{\theta}_j, j = 1, \dots, T)$. Para minimizar

$$Q(\boldsymbol{\theta}) = \sum_{s=0}^{N-1} \sum_{j=1}^T \left(X_{j+sT} - \alpha_j^{(1)} X_{j+sT-1} I_j^{(1)} + \alpha_j^{(2)} X_{j+sT-1} I_j^{(2)} + \lambda_j \right)^2,$$

é necessária a resolução de T sistemas da forma:

$$\begin{cases} \frac{\partial Q}{\partial \alpha_j^{(1)}} = 0 \\ \frac{\partial Q}{\partial \alpha_j^{(2)}} = 0 \\ \frac{\partial Q}{\partial \lambda_j} = 0 \end{cases}, j = 1, \dots, T. \quad (3.4)$$

A representação matricial do sistema (3.4), relativamente aos parâmetros associados ao período j , é dado por

$$\begin{bmatrix} \sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(1)} & 0 & \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} \\ 0 & \sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(2)} & \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} \\ \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} & \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} & N \end{bmatrix} \begin{bmatrix} \alpha_j^{(1)} \\ \alpha_j^{(2)} \\ \lambda_j \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(1)} \\ \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(2)} \\ \sum_{s=0}^{N-1} X_{j+sT} \end{bmatrix}.$$

Assim

$$\begin{aligned} \sum_{s=0}^{N-1} \alpha_j^{(1)} X_{j+sT-1}^2 I_j^{(1)} + \sum_{s=0}^{N-1} \lambda_j X_{j+sT-1} I_j^{(1)} &= \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(1)} \\ \alpha_j^{(1)} \sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(1)} &= \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(1)} - \lambda_j \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} \\ \alpha_j^{(1)} &= \frac{\sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(1)} - \lambda_j \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)}}{\sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(1)}} \end{aligned} \quad (3.5)$$

$$\begin{aligned} \sum_{s=0}^{N-1} \alpha_j^{(2)} X_{j+sT-1}^2 I_j^{(2)} + \sum_{s=0}^{N-1} \lambda_j X_{j+sT-1} I_j^{(2)} &= \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(2)} \\ \alpha_j^{(2)} \sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(2)} &= \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(2)} - \lambda_j \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} \\ \alpha_j^{(2)} &= \frac{\sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(2)} - \lambda_j \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)}}{\sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(2)}} \end{aligned} \quad (3.6)$$

$$\begin{aligned} \sum_{s=0}^{N-1} \alpha_j^{(1)} X_{j+sT-1} I_j^{(1)} + \sum_{s=0}^{N-1} \alpha_j^{(2)} X_{j+sT-1} I_j^{(2)} + \sum_{s=0}^{N-1} \lambda_j &= \sum_{s=0}^{N-1} X_{j+sT} \\ \sum_{s=0}^{N-1} \lambda_j &= \sum_{s=0}^{N-1} X_{j+sT} - \sum_{s=0}^{N-1} \alpha_j^{(1)} X_{j+sT-1} I_j^{(1)} - \sum_{s=0}^{N-1} \alpha_j^{(2)} X_{j+sT-1} I_j^{(2)} \\ N \lambda_j &= \sum_{s=0}^{N-1} X_{j+sT} - \alpha_j^{(1)} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} - \alpha_j^{(2)} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} \end{aligned}$$

$$\lambda_j = N^{-1} \left(\sum_{s=0}^{N-1} X_{j+sT} - \alpha_j^{(1)} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} - \alpha_j^{(2)} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} \right).$$

Desta forma os estimadores dos MQC para o modelo PSETINAR são dados por

$$\left\{ \begin{array}{l} \hat{\alpha}_{j,MQC}^{(1)} = \frac{N \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(1)} - \sum_{s=0}^{N-1} X_{j+sT} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)}}{N \sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(1)} - \left(\sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} \right)^2} \\ \hat{\alpha}_{j,MQC}^{(2)} = \frac{N \sum_{s=0}^{N-1} X_{j+sT} X_{j+sT-1} I_j^{(2)} - \sum_{s=0}^{N-1} X_{j+sT} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)}}{\left(N \sum_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(2)} - \left(\sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} \right)^2 \right)} \\ \hat{\lambda}_{j,MQC} = N^{-1} \left(\sum_{s=0}^{N-1} X_{j+sT} - \alpha_j^{(1)} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(1)} - \alpha_j^{(2)} \sum_{s=0}^{N-1} X_{j+sT-1} I_j^{(2)} \right) \end{array} \right. , \quad j = 1, \dots, T.$$

A distribuição assintótica dos estimadores MQC pode ser obtida usando os resultados de Klimko & Nelson (1978).

Teorema 3.1. *Num modelo PSETINAR(2; 1, 1)_T os estimadores de MQC são assitoticamente normais, i.e.,*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, \mathbf{V}^{-1} \mathbf{W} \mathbf{V}^{-1}),$$

onde \mathbf{V} e \mathbf{W} são matrizes quadradas de ordem $3T$ definidas por blocos de dimensão 3×3 dadas por,

$$\mathbf{V} = \begin{bmatrix} \boldsymbol{\Psi}_{:1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{:2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Psi}_{:T} \end{bmatrix} \quad e \quad \mathbf{W} = \begin{bmatrix} \boldsymbol{\Omega}_{:1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{:2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Omega}_{:T} \end{bmatrix},$$

onde cada bloco Ψ_j e Ω_j tem seus elementos definidos, respetivamente, por

$$\Psi_{(k,l):j} = E \left[\frac{\partial}{\partial \theta_{k:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \frac{\partial}{\partial \theta_{l:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \right],$$

$$\Omega_{(k,l):j} = E \left[U_{j+sT}^2 \frac{\partial}{\partial \theta_{a:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \frac{\partial}{\partial \theta_{l:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \right],$$

com $k, l = 1, 2, 3$ e $j = 1, \dots, T$, e $\boldsymbol{\theta}_j := (\theta_{1:j}, \theta_{2:j}, \theta_{3:j}) \equiv (\alpha_j^{(1)}, \alpha_j^{(2)}, \lambda_j)$, são os parâmetros associados ao período j .

Demonstração. Monteiro *et al.* (2010, 2012) verificaram as condições de normalidade assintótica de Klimko & Nelson (1978) para modelos PINAR(1) $_T$ e modelos SETINAR(2; 1, 1) respetivamente, onde a matriz de covariância assintótica é dada por $\mathbf{J}^{-1} = \mathbf{V}^{-1} \mathbf{W} \mathbf{V}^{-1}$. Antes de obter as matrizes \mathbf{V} e \mathbf{W} para o modelo PSETINAR(2; 1, 1) $_T$ define-se:

$$\begin{aligned} u_j^{(1)} &:= E[X_{j-1+sT} | X_{j-1+sT-d} \leq r_j], & u_j^{(2)} &:= E[X_{j-1+sT} | X_{j-1+sT-d} > r_j], \\ m_{j,i}^{(1)} &:= E[X_{j-1+sT}^i | X_{j-1+sT-d} \leq r_j], & m_{j,i}^{(2)} &:= E[X_{j-1+sT}^i | X_{j-1+sT-d} > r_j], & q_j^{(1)} &:= \\ &P[X_{j+sT-d} \leq r_j], & & & q_j^{(1)} &:= P[X_{j+sT-d} > r_j]. \end{aligned}$$

As componentes das matrizes \mathbf{V} e \mathbf{W} , são matrizes por blocos de dimensão 3×3 tais que

$$\mathbf{V} = \begin{bmatrix} \Psi_{:1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Psi_{:2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Psi_{:T} \end{bmatrix} \quad e \quad \mathbf{W} = \begin{bmatrix} \Omega_{:1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Omega_{:2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Omega_{:T} \end{bmatrix},$$

onde cada bloco $\Psi_{:j}$ e $\Omega_{:j}$ é dado por

$$\Psi_{:j} = \begin{bmatrix} q_j^{(1)} m_{j,2}^{(1)} & 0 & q_j^{(1)} u_j^{(1)} \\ 0 & q_j^{(2)} m_{j,2}^{(2)} & q_j^{(2)} u_j^{(2)} \\ q_j^{(1)} u_j^{(1)} & q_j^{(2)} u_j^{(2)} & 1 \end{bmatrix}$$

e

$$\Omega_{:j} = \begin{bmatrix} q_j^{(1)}(\alpha_j^{(1)}(1 - \alpha_j^{(1)})m_{j,3}^{(1)} + \sigma_j^2 m_{j,2}^{(1)}) & 0 & q_j^{(1)}(\alpha_j^{(1)}(1 - \alpha_j^{(1)})m_{j,2}^{(1)} + \sigma_j^2 u_j^{(1)}) \\ 0 & q_j^{(2)}\alpha_j^{(2)}(1 - \alpha_j^{(2)})m_{j,3}^{(2)} + \sigma_j^2 m_{j,2}^{(2)} & q_j^{(1)}(\alpha_j^{(2)}(1 - \alpha_j^{(2)})m_{j,2}^{(2)} + \sigma_j^2 u_j^{(2)}) \\ q_j^{(1)}(\alpha_j^{(1)}(1 - \alpha_j^{(1)})m_{j,2}^{(1)} + \sigma_j^2 u_j^{(1)}) & q_j^{(2)}(\alpha_j^{(2)}(1 - \alpha_j^{(2)})m_{j,2}^{(2)} + \sigma_j^2 u_j^{(2)}) & \sigma_j^2(q_j^{(1)}m_{j,2}^{(1)} + q_j^{(2)}m_{j,2}^{(2)} + 1) \end{bmatrix}.$$

Calculando $\mathbf{J}^{-1} = \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}$ a demonstração fica completa desde que a matriz \mathbf{V} seja não singular. Para verificar essa condição basta verificar que o bloco genérico $\Psi_{:j}$ é não singular, ou seja, $\det|\Psi_{:j}| \neq 0, \forall j = 1, \dots, T$. Dado que $\Psi_{:j}$ é definida como sendo uma matriz não negativa, isto é o mesmo que assumir que $\Psi_{:j}$ é não singular o que completa a prova. ■

3.2.2 Estimadores de Máxima Verosimilhança Condicional

Dado x_0 fixo, a função de verosimilhança condicional, para o modelo PSETINAR(2; 1, 1) $_T$, é dada por

$$\begin{aligned} L(\boldsymbol{\theta}) &:= \prod_{s=0}^{N-1} \prod_{j=1}^T P_j(X_{j+sT} = x_{j+sT} | X_{j-1+sT} = x_{j-1+sT}) \\ &= \prod_{s=0}^{N-1} \prod_{j=1}^T p_j(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j) \end{aligned}$$

com

$$I_j^{(1)} := \begin{cases} 1, & X_{j+sT-d} \leq r_j \\ 0, & X_{j+sT-d} > r_j \end{cases} \quad e \quad I_j^{(2)} = 1 - I_j^{(1)}$$

e

$$\begin{aligned} p_j(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j) &= \\ &= e^{-\lambda_j} \sum_{m=0}^{\min(x_{j-1+sT}, x_{j+sT})} \sum_{k=1}^2 C_m^{x_{j-1+sT}} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{(x_{j-1+sT}-m)} \frac{\lambda^{x_{j+sT}-m}}{(x_{j+sT}-m)!} I_j^{(k)} \\ &= p_j(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)}, \lambda_j) + p_j(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(2)} I_j^{(2)}, \lambda_j) \\ &=: p_j(x_{j-1+sT}, x_{j+sT})^{(1)} + p_j(x_{j-1+sT}, x_{j+sT})^{(2)}. \end{aligned}$$

Os estimadores **MVC**,

$$\hat{\boldsymbol{\theta}}_{MVC} := \left(\hat{\alpha}_1^{(1)}{}_{MVC}, \hat{\alpha}_1^{(2)}{}_{MVC}, \hat{\lambda}_{1MVC}, \dots, \hat{\alpha}_T^{(1)}{}_{MVC}, \hat{\alpha}_T^{(2)}{}_{MVC}, \hat{\lambda}_{TMVC} \right),$$

são os valores que maximizam o logaritmo da função de máxima verosimilhança condicional

$$l(\boldsymbol{\theta}) = \sum_{s=0}^{N-1} \sum_{j=1}^T \ln p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right).$$

Para a obtenção dos estimadores **MVC** calcula-se as derivadas parciais de primeira ordem de $p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)$ em relação aos parâmetros $\alpha_j^{(1)}$, $\alpha_j^{(2)}$ e λ_j com $j = 1, \dots, T$. Por simplicidade nas equações que seguem fazem-se as seguintes substituições:

$$x_{j-1+sT} = a, \quad x_{j+sT} = b \quad e \quad p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right) = p_j(b|a).$$

Para cada ramo $k = 1, 2$, a derivada parcial da função de verosimilhança relativo aos parâmetros $\alpha_j^{(k)}$ é dada por

$$\begin{aligned}
 \frac{\partial p_j(b|a)}{\partial \alpha_j^{(k)}} &= \frac{\partial}{\partial \alpha_j^{(k)}} \left[e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} C_m^a \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \frac{\lambda_j^{b-m}}{(b-m)!} \right] \\
 &= e^{-\lambda_j} I_j^{(k)} \sum_{m=0}^{\min(a,b)} \left[C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \left(m \alpha_j^{(k)m-1} (1 - \alpha_j^{(k)})^{a-m} - \right. \right. \\
 &\quad \left. \left. - b^m (a-m) (1 - \alpha_j^{(k)})^{a-m-1} \right) \right] \\
 &= e^{-\lambda_j} I_j^{(k)} \sum_{m=0}^{\min(a,b)} \left[C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \left(\frac{m}{\alpha_j^{(k)}} - \frac{a-m}{1 - \alpha_j^{(k)}} \right) \right] \\
 &= \frac{e^{-\lambda_j}}{\alpha_j^{(k)} (1 - \alpha_j^{(k)})} I_j^{(k)} \sum_{m=0}^{\min(a,b)} \left[C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{b-m} (m - a \alpha_j^{(k)}) \right] \\
 &= \frac{e^{-\lambda_j}}{\alpha_j^{(k)} (1 - \alpha_j^{(k)})} I_j^{(k)} \sum_{m=0}^{\min(a,b)} \left[m C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \right. \\
 &\quad \left. - a \alpha_j^{(k)} C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \right] \\
 &= \frac{I_j^{(k)}}{\alpha_j^{(k)} (1 - \alpha_j^{(k)})} \left[\underbrace{e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} \left(m C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \right)}_A \right. \\
 &\quad \left. - e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} \left(a \alpha_j^{(k)} C_m^a \frac{\lambda_j^{b-m}}{(b-m)!} \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \right) \right]. \tag{3.7}
 \end{aligned}$$

Para o termo A na equação (3.7), quer-se mostrar que

$$\begin{aligned}
 A &= e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} m C_m^a \alpha_j^{(k)m} (1 - \alpha_j^{(k)})^{a-m} \frac{\lambda_j^{b-m}}{(b-m)!} \\
 &= m p_j(b|a)^{(k)} = b p_j(b|a)^{(k)} - \lambda_j p_j(b-1|a)^{(k)}.
 \end{aligned}$$

Assim,

$$\begin{aligned}
 e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} m C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-m}}{(b-m)!} &= m p_j(b|a)^{(k)} + b p_j(b|a)^{(k)} - b p_j(b|a)^{(k)} \\
 &= (m-b) p_j(b|a)^{(k)} + b p_j(b|a)^{(k)} \\
 &= -(b-m) e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} m C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-m}}{(b-m)!} + b p_j(b|a)^{(k)} \\
 &= -e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} m C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-m}}{(b-1-m)!} \frac{(b-m)}{(b-m)} + b p_j(b|a)^{(k)} \\
 &= -e^{-\lambda_j} \lambda_j \sum_{m=0}^{\min(a,b)} m C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-1-m}}{(b-1-m)!} + b p_j(b|a)^{(k)} \\
 &= -\lambda_j p_j(b-1|a)^{(k)} + b p_j(b|a)^{(k)}.
 \end{aligned} \tag{3.8}$$

Substituindo (3.8) em (3.7) tem-se a derivada parcial da função de verosimilhança relativo aos parâmetros $\alpha_j^{(k)}$

$$\begin{aligned}
 \frac{\partial p_j(b|a)}{\partial \alpha_j^{(k)}} &= \frac{I_j^{(k)}}{\alpha_j^{(k)} \left(1 - \alpha_j^{(k)}\right)} \left[b p_j(b|a)^{(k)} - \lambda_j p_j(b-1|a)^{(k)} - a \alpha_j^{(k)} p_j(b|a)^{(k)} \right] \\
 &= \frac{I_j^{(k)}}{\alpha_j^{(k)} \left(1 - \alpha_j^{(k)}\right)} \left[p_j(b|a)^{(k)} (b - a \alpha_j^{(k)}) - \lambda_j p_j(b-1|a)^{(k)} \right].
 \end{aligned}$$

A derivada parcial de $p_j^{(k)}$ em relação a λ_j é dada por

$$\frac{\partial p_j(b|a)^{(k)}}{\partial \lambda_j} = \frac{\partial p_j(b|a)^{(1)}}{\partial \lambda_j} + \frac{\partial p_j(b|a)^{(2)}}{\partial \lambda_j},$$

onde para cada ramo $k = 1, 2$, tem-se

$$\frac{\partial p_j(b|a)^{(k)}}{\partial \lambda_j} = \frac{\partial}{\partial \lambda_j} \left[e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-m}}{(b-m)!} \right]$$

$$\begin{aligned}
 &= e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-1-m}(b-m)}{(b-m)!} - \\
 &- e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-m}}{(b-m)!} \\
 &= e^{-\lambda_j} \sum_{m=0}^{\min(a,b)} C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-1-m}(b-m)}{(b-m)!} - p_j(b|a)^{(k)} \\
 &= e^{-\lambda_j} \sum_{m=0}^{\min(a-1,b)} C_m^a \alpha_j^{(k)m} \left(1 - \alpha_j^{(k)}\right)^{a-m} \frac{\lambda_j^{b-1-m}(b-m)}{(b-m)!} - p_j(b|a)^{(k)} \\
 &= p_j(b-1|a)^{(k)} - p_j(b|a)^{(k)}.
 \end{aligned}$$

As derivadas de primeira ordem da função de verosimilhança são dadas por:

$$\left\{ \begin{aligned} \frac{\partial l(\theta)}{\partial \alpha_j^{(k)}} &= \frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_{s=0}^{N-1} (b - a\alpha_j^{(k)}) - \lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \\ \frac{\partial l(\theta)}{\partial \lambda_j} &= \sum_{s=0}^{N-1} \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} - N \end{aligned} \right. ,$$

para $k = 1, 2$ e $j = 1, \dots, T$. Igualando a zero resulta em

$$\left\{ \begin{aligned} \frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_{s=0}^{N-1} (b - a\alpha_j^{(k)}) - \lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} I_j^{(k)} &= 0 \\ \sum_{s=0}^{N-1} \left(\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)} I_j^{(1)} + \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} I_j^{(2)} \right) &= N \end{aligned} \right. ,$$

para $k = 1, 2$ e $j = 1, \dots, T$. Substituindo a , b e $p_j(b|a)$ pelos seus valores originais tem-se os estimadores de máxima verosimilhança condicional dados pela resolução de

T sistemas da forma

$$\left\{ \begin{array}{l} \frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_{s=0}^{N-1} (x_{j+sT} - x_{j-1+sT} \alpha_j^{(k)}) - \\ - \lambda_j \frac{p_j(x_{j-1+sT}, x_{j+sT} - 1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j)}{p_j(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j)} = 0 \\ \sum_{s=0}^{N-1} \frac{p_j(x_{j-1+sT}, x_{j+sT} - 1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j)}{p_j(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j)} = N \end{array} \right.,$$

para $k = 1, 2$ e $j = 1, \dots, T$. O sistema anterior não tem solução analítica sendo, portanto, necessário usar métodos numéricos para o resolver. No entanto, para λ_j , tem-se que no entanto resulta

$$\hat{\lambda}_j = \frac{1}{N} \sum_{s=0}^{N-1} (x_{j+sT} - \hat{\alpha}_j^{(k)} x_{j-1+sT}).$$

A distribuição assintótica dos estimadores de máxima verosimilhança condicional é apresentada no próximo resultado, cálculos auxiliares podem ser encontrados nos anexos.

Teorema 3.2. *Para o modelo $PSETINAR(2; 1, 1)_T$ os estimadores de máxima verosimilhança tem a seguinte distribuição assintótica*

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}_1^{(1)} - \alpha_1^{(1)} \\ \hat{\alpha}_1^{(2)} - \alpha_1^{(2)} \\ \hat{\lambda}_1 - \lambda_1 \\ \vdots \\ \hat{\alpha}_T^{(1)} - \alpha_T^{(1)} \\ \hat{\alpha}_T^{(2)} - \alpha_T^{(2)} \\ \hat{\lambda}_T - \lambda_T \end{pmatrix} \xrightarrow{D} N(\mathbf{0}, \mathbf{I}^{-1}), \text{ onde } \mathbf{I} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{M}_T \end{bmatrix},$$

é a matriz de informação de Fisher e, para $j = 1, \dots, T$,

$$\mathbf{M}_j = \begin{bmatrix} -E \left[\frac{\partial^2 l(\theta)}{\partial (\alpha_j^{(1)})^2} \right] & -E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(1)} \partial \alpha_j^{(2)}} \right] & -E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(1)} \partial \lambda_j} \right] \\ -E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(1)} \partial \alpha_j^{(2)}} \right] & -E \left[\frac{\partial^2 l(\theta)}{\partial (\alpha_j^{(2)})^2} \right] & -E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(2)} \partial \lambda_j} \right] \\ -E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(1)} \partial \lambda_j} \right] & -E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(2)} \partial \lambda_j} \right] & -E \left[\frac{\partial^2 l(\theta)}{\partial \lambda_j^2} \right] \end{bmatrix},$$

sendo

$$\begin{aligned} E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)2}} \right] &= N \sum_a^{+\infty} \sum_b^{+\infty} P(X_{j+sT} = b) \frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \left\{ \left[(2\alpha_j^{(k)} - 1)b - a\alpha_j^{(k)2} \right] \right. \\ &\quad \left. p_j(b|a) + 2(1 - \alpha_j^{(k)})\lambda_j p_j(b - 1|a)^{(k)} + \lambda_j^2 p_j(b - 2|a)^{(k)} - \lambda_j^2 \frac{p_j^2(b - 1|a)^{(k)}}{p_j(b|a)^{(k)}} \right\} \\ E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)} \partial \lambda_j} \right] &= \frac{N}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_a^{+\infty} \sum_b^{+\infty} P(X_{j+sT} = b) \left\{ -p_j(b - 1|a)^{(k)} - \lambda_j p_j(b - 2|a)^{(k)} \right. \\ &\quad \left. + \lambda_j \frac{p_j^2(b - 1|a)^{(k)}}{p_j(b|a)^{(k)}} \right\} \\ E \left[\frac{\partial^2 l(\theta)}{\partial \lambda_j^2} \right] &= N \sum_a^{+\infty} \sum_b^{+\infty} P(X_{j+sT} = b) \left\{ \lambda_j p_j(b - 2|a)^{(k)} - \frac{p_j^2(b - 1|a)^{(k)}}{p_j(b|a)^{(k)}} \right\}. \end{aligned}$$

Demonstração. Franke & Seligmann (1993) mostraram que as 6 condições a seguir são verificadas para um modelo quando as inovações tem distribuição de Poisson, que é o caso do modelo PSETINAR. Verificaram ainda que estas condições implicam as condições dos resultados de Billingsley (1961) sobre os estimadores de um processo de Markov para um modelo INAR.

- [C1] O conjunto $z = P(Z_{j+sT} = z) = f(z, \lambda_j) = e^{-\lambda_j} \frac{\lambda_j^z}{z!}$ não depende de λ_j ;
- [C2] $E[Z_{j+sT}^3] < \infty$;
- [C3] $P(Z_{j+sT} = z)$ é 3 vezes diferenciável relativamente a λ_j no espaço dos parâmetros, B ;

- [C4] Para todo $\lambda_j \in B$ existe uma vizinhança U de λ_j tal que

1. $\sum_{z=0}^{\infty} \sup_{\lambda_j \in U} f(z, \lambda_j) < \infty;$
2. $\sum_{z=0}^{\infty} \sup_{\lambda_j \in U} \left| \frac{\partial f(z, \lambda_j)}{\partial \lambda_j} \right| < \infty;$
3. $\sum_{z=0}^{\infty} \sup_{\lambda_j \in U} \left| \frac{\partial^2 f(z, \lambda_j)}{\partial \lambda_j^2} \right| < \infty.$

- [C5] Para qualquer $\lambda_j \in B$ existe uma vizinhança U de λ_j e sequências crescentes $\psi_j(n) = \text{const}1.n$, $\psi_{ji}(n) = \text{const}2.n^2$ e $\psi_{jil}(n) = \text{const}3.n^3$, $\text{const}1$, $\text{const}2$ e $\text{const}3$ constantes adequadas, $n \geq 0$, tal que $\forall \lambda_j \in B$ e $\forall m \leq n$ com $f(z, \lambda_j)$ não nulos:

$$\begin{aligned} * \quad & \left| \frac{\partial f(m, \lambda_j)}{\partial \lambda_j} \right| \leq \psi_j(n) f(m, \lambda_j); \\ * \quad & \left| \frac{\partial^2 f(m, \lambda_j)}{\partial \lambda_j^2} \right| \leq \psi_{ji}(n) f(m, \lambda_j); \\ * \quad & \left| \frac{\partial^3 f(m, \lambda_j)}{\partial \lambda_j^3} \right| \leq \psi_{jil}(n) f(m, \lambda_j) \end{aligned}$$

e relativamente à estacionariedade de (X_{j+sT})

$$\begin{aligned} * \quad & E [\psi_j^3(X_{j+sT+1})] < \infty; \\ * \quad & E [X_{j+sT} \psi_{ji}(X_{j+sT+1})] < \infty \\ * \quad & E [\psi_j(X_{j+sT}) \psi_{ji}(X_{j+sT+1})] < \infty; \\ * \quad & E [\psi_{jil}(X_{j+sT})] < \infty. \end{aligned}$$

- [C6] A matriz de informação de Fisher $\sum(\theta)$ é não singular.

As condições [C1]-[C5] dizem respeito somente a distribuição das inovações e, como tal são, Monteiro (2010) e Monteiro *et al.* (2012) mostraram que tais implicam as condições dos resultados de Billingsley (1961) sobre os estimadores dos parâmetros de processos de Markov também para os modelos PINAR(1)_T e SETINAR(2; 1, 1) respetivamente. Dado que para cada um dos períodos do modelo PSETINAR(2; 1, 1)_T, $p_j(\cdot|\cdot)$ é a função de transição do modelo SETINAR com parâmetros $\alpha_j^{(1)}$, $\alpha_j^{(2)}$ e λ_j ,

fica garantido que também neste caso as condições [C1]-[C6] impliquem nas condições gerais do Teorema de Billingsley (1961) para o modelo $\text{PSETINAR}(2; 1, 1)_T$. ■

3.3 Estudo de Simulação

O objetivo do estudo de simulação apresentado nesta secção é ilustrar o comportamento dos estimadores MQC, e MVC apresentados em secções anteriores e aplicados à amostras de dimensões variadas. Foram simuladas 1000 realizações do processo definido na Secção 3.1, com período $T = 2$, dado por

$$X_t = \phi_t \circ X_{t-1} + Z_t.$$

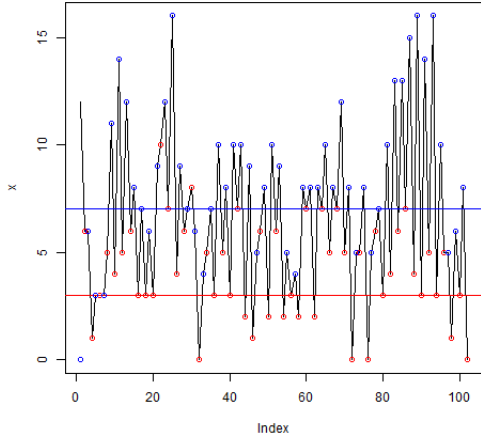
As amostras utilizadas apresentam dimensão $n = 2N$, com $N = 100, 250, 500$ e o estudo de simulação contempla 4 diferentes conjuntos de parâmetros $\theta = (\alpha^{(1)}, \alpha^{(2)}, \lambda^{(1)}, \lambda^{(2)}, d, r)$ apresentados na Tabela 3.1. Os valores escolhidos para os parâmetros

Tabela 3.1: Parâmetros dos modelos $\text{PSETINAR}(2; p^{(1)}, p^{(2)})$

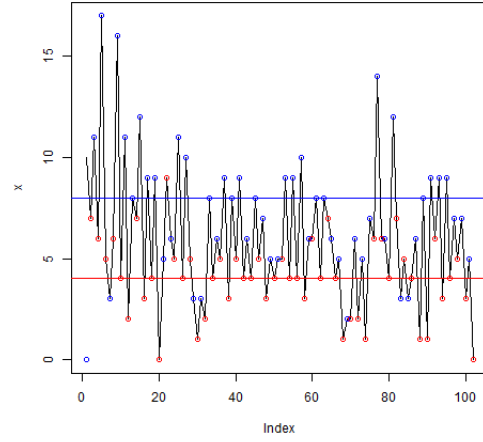
	$\alpha_1^{(1)}$	$\alpha_2^{(1)}$	$\alpha_1^{(2)}$	$\alpha_2^{(2)}$	$\lambda^{(1)}$	$\lambda^{(2)}$	$r^{(1)}$	$r^{(2)}$
M_1	0.2	0.65	0.2	0.1	7	3	8	4
M_2	0.2	0.1	0.2	0.1	7	3	8	4
M_3	0.6	0.3	0.6	0.35	6	6	9	12
M_4	0.8	0.1	0.65	0.2	7	3	8	4

tem em conta modelos cujos os valores de $\alpha_1^{(k)}$ são todos inferiores a 0.5 ou todos superiores a 0.5. Os modelos M_1 e M_2 diferem somente no valor do parâmetro $\alpha_2^{(1)}$, sendo a principal característica nestes dois modelos que os valores dos parâmetros α são não superiores a 0.5 excetuando-se no modelo M_1 onde $\alpha_2^{(1)} = 0.65$. Já os modelos M_3 e M_4 apresentam os valores dos parâmetros $\alpha_1^{(k)}$ maiores que 0.5 e todos os parâmetros $\alpha_2^{(k)}$ apresentam valores inferiores a 0.5.

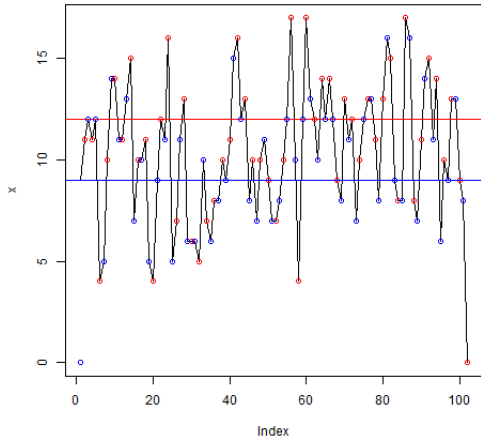
A Figura 3.1 apresenta os gráficos de algumas realizações dos modelos M_1 à M_4 . Em cada um dos gráficos tem-se 2 linhas horizontais que representam os limiares r de cada um dos períodos.



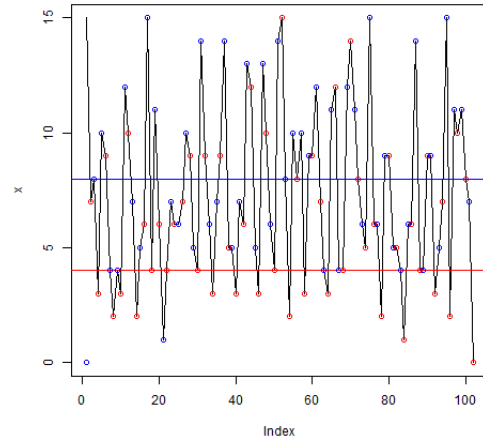
(a) $M_1 : \lambda^{(1)} \neq \lambda^{(2)}$



(b) $M_2 : \lambda^{(1)} = \lambda^{(2)}$



(c) $M_3 : \lambda^{(1)} = \lambda^{(2)}$



(d) $M_4 : \lambda^{(1)} \neq \lambda^{(2)}$

Figura 3.1: Gráfico de algumas realizações do modelo PSETINAR.

Para cada conjunto de parâmetros o valor do limiar r foi escolhido de forma que, para cada réplica, entre 30% e 70% das observações estivessem num dos ramos. O parâmetro de atraso, d , foi considerado constante igual a 1 para todos os conjuntos de dados. Foram simuladas 2000 conjuntos de dados independentes para cada um dos modelos e exigiu-se sempre que os valores estimados estivessem dentro do suporte dos parâmetros. Desta forma o número final de réplicas admitidas são 500 de dimensão 1000 para cada um dos modelos. A análise foi baseada nas n primeiras observações para $n = 2N = 200, 500, 1000$, sendo N a dimensão de cada período T .

A apresentação e discussão dos resultados consiste num apanhado geral dos dois métodos de estimação, seus valores estimados e análises gráficas.

3.3.1 Análise dos Resultados da Estimação

As médias amostrais obtidas para os parâmetros estimados $\hat{\theta}$ dos modelos M_1 e M_2 são apresentados na Tabela 3.2 e o enviesamento de tais estimativas são apresentadas nas Figuras 3.2 e 3.3. Para os modelos M_3 e M_4 , as médias amostrais dos parâmetros estimados $\hat{\theta}$ são apresentados na Tabela 3.3 e o enviesamento de tais estimativas podem ser verificados nas Figuras 3.4 e 3.5. A partir das tabela e figuras referidas, conclui-se que para o modelo PSETINAR, ambos os métodos de estimação utilizados, o método dos mínimos quadrados e método da máxima verosimilhança condicional, apresentaram resultados bastante semelhantes. Cabe destacar que as estimativas obtidas pelos métodos MQC apresentam maior dispersão em relação à média que as estimativas obtidas pelo MV. Para todos os parâmetros, tanto para as amostras mais pequenas ($n = 200$) como para as amostras de grande dimensão ($n = 1000$), o método MV apresenta menor viés nas estimativas.

Considerem-se os casos em que as realizações são de pequena e média dimensão. Para cada período, as estimativas de $\alpha_j^{(k)}$, ($k = 1, 2$), apresentam o mesmo comportamento encontrado nas estimativas do modelo SETINAR($2; p^{(1)}, p^{(2)}$). Quando α é inferior a 0.5 as respectivas estimativas são mais dispersas e possuem assimetria positiva. Quando α é maior que 0.5 as estimativas apresentam maior simetria. Em todos os modelos as estimativas de $\lambda^{(1)}$ são subestimadas e apresentam um maior viés quando comparadas com as estimativas do parâmetro $\lambda^{(2)}$. Globalmente pode dizer-se que: não existe diferença significativa entre as estimativas de MQC e de MV; existe tendência para sobrestimação dos parâmetros autorregressivos; relativamente ao desempenho dos estimadores dos parâmetros de inovação nota-se, de uma forma geral, que subestimam os parâmetros e têm um desempenho pior, apesar de não muito evidente, quando existe uma maior variabilidade no modelo.

Tabela 3.2: Média amostral das estimativas de θ para os modelos M_1 e M_2

Modelo	θ_i	n					
		200		500		1000	
		MQC	MV	MQC	MV	MQC	MV
M_1	$\alpha_1^{(1)}$	0.2872	0.2788	0.2362	0.2299	0.2230	0.2188
	0.2	(0.0872)	(0.0788)	(0.0362)	(0.0299)	(0.0230)	(0.0188)
	$\alpha_2^{(1)}$	0.7154	0.7051	0.6811	0.6712	0.6701	0.6637
	0.65	(0.0654)	(0.0551)	(0.0311)	(0.0212)	(0.0201)	(0.0137)
	$\lambda^{(1)}$	6.6280	6.6360	6.8333	6.8590	6.8990	6.9260
	7	(0.3720)	(0.3640)	(0.1670)	(0.1410)	(0.1010)	(0.0740)
	$\alpha_1^{(2)}$	0.2189	0.2206	0.2074	0.2088	0.2040	0.2024
	0.2	(0.0189)	(0.0206)	(0.0074)	(0.0088)	(0.0040)	(0.0024)
	$\alpha_2^{(2)}$	0.1153	0.1128	0.1060	0.1062	0.1032	0.1020
	0.1	(0.0153)	(0.0128)	(0.0060)	(0.0062)	(0.0032)	(0.0020)
	$\lambda^{(2)}$	2.8380	2.8310	2.9320	2.9250	2.959	2.9750
	3	(0.1620)	(0.1690)	(0.0680)	(0.0075)	(0.0410)	(0.0025)
M_2	$\alpha_1^{(1)}$	0.2842	0.2760	0.2337	0.2298	0.2163	0.2200
	0.2	(0.0842)	(0.0760)	(0.0337)	(0.0298)	(0.0163)	(0.0200)
	$\alpha_2^{(1)}$	0.1810	0.1747	0.1322	0.1324	0.1156	0.1193
	0.1	(0.0810)	(0.0747)	(0.0322)	(0.0394)	(0.0156)	(0.0193)
	$\lambda^{(1)}$	6.6100	6.6450	6.8330	6.8570	6.9210	6.9070
	7	(0.3900)	(0.3550)	(0.1670)	(0.1430)	(0.0790)	(0.0930)
	$\alpha_1^{(2)}$	0.2198	0.2164	0.2109	0.2053	0.2071	0.2000
	0.2	(0.0198)	(0.0164)	(0.0109)	(0.0053)	(0.0071)	(0.0000)
	$\alpha_2^{(2)}$	0.1168	0.1174	0.1087	0.1044	0.1055	0.0994
	0.1	(0.0168)	(0.0174)	(0.0087)	(0.0044)	(0.0055)	(0.0006)
	$\lambda^{(2)}$	2.8400	2.8690	2.9030	2.9520	2.9410	2.9910
	3	(0.1600)	(0.1310)	(0.0970)	(0.0480)	(0.0590)	(0.0090)

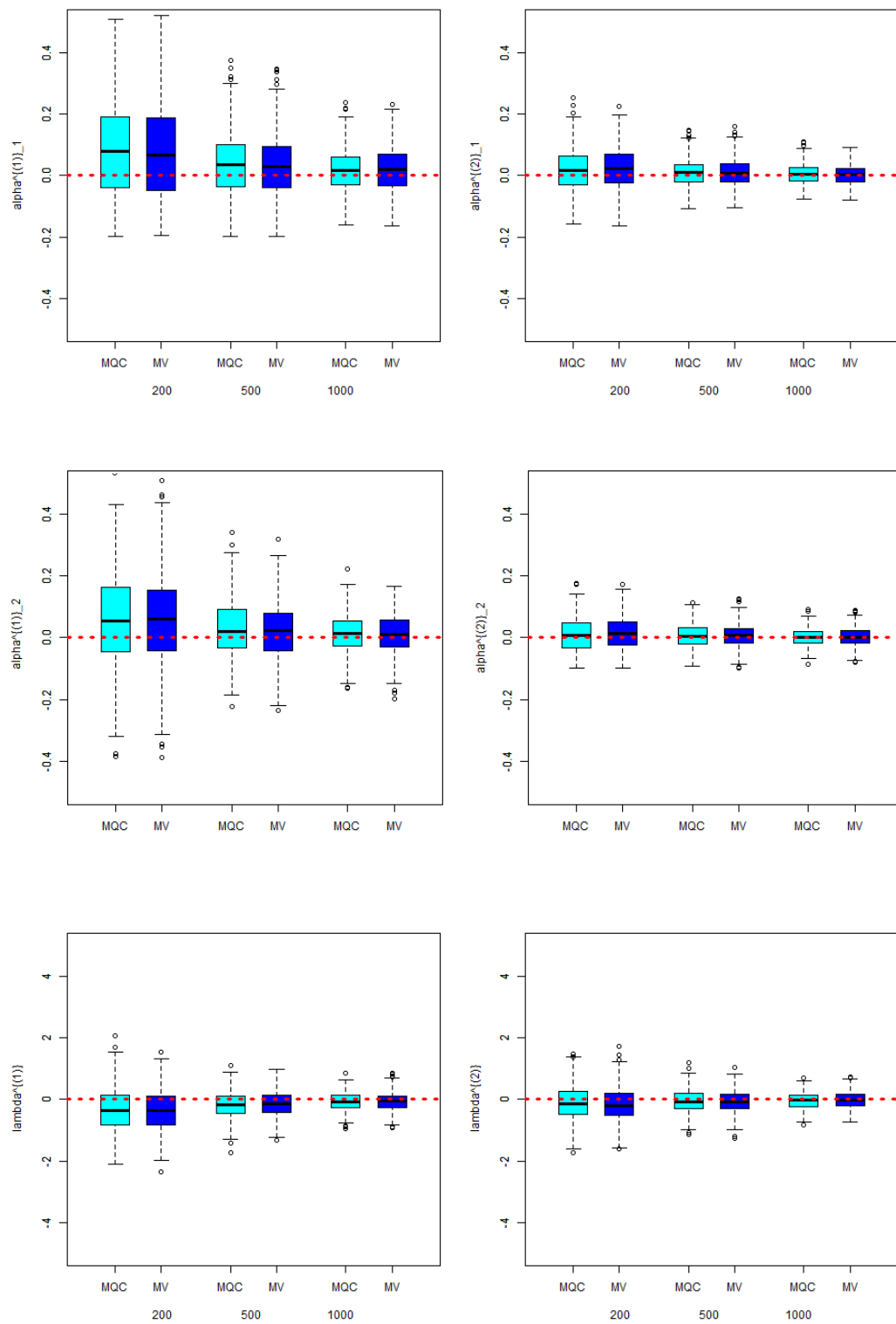


Figura 3.2: Viés das estimativas de $\theta = (0.2, 0.65; 0.2, 0.1; 7; 3)$ do modelo M_1 .

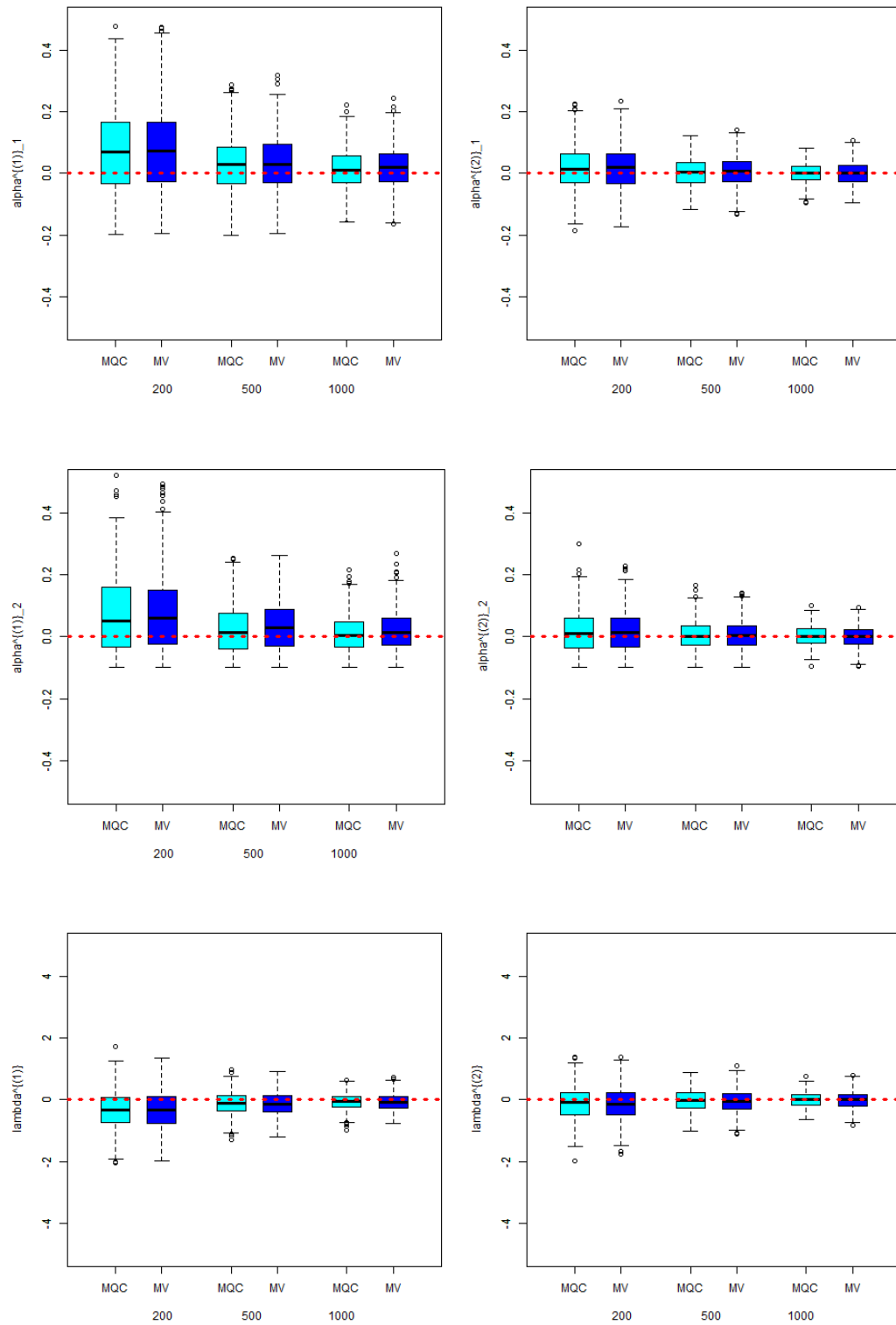


Figura 3.3: Viés das estimativas de $\theta = (0.2, 0.1; 0.2, 0.1; 7; 3)$ do modelo M_2 .

Tabela 3.3: Média amostral das estimativas de θ para os modelos M_3 e M_4 .

Modelo	θ_i	n					
		200		500		1000	
		MQC	MV	MQC	MV	MQC	MV
M_3	$\alpha_1^{(1)}$	0.6760	0.6636	0.6318	0.6291	0.6156	0.6149
	0.65	(0.0260)	(0.0136)	(0.0182)	(0.0209)	(0.0344)	(0.0311)
	$\alpha_2^{(1)}$	0.3584	0.3529	0.3241	0.3243	0.3129	0.3127
	0.3	(0.0584)	(0.0529)	(0.0241)	(0.0243)	(0.0129)	(0.0127)
	$\lambda^{(1)}$	5.1550	5.3130	5.6500	5.6940	5.8180	5.8270
	6	(0.0845)	(0.6870)	(0.0350)	(0.3060)	(0.0182)	(0.1730)
	$\alpha_1^{(2)}$	0.6026	0.6040	0.5979	0.6046	0.59990	0.6026
	0.60	(0.0026)	(0.0040)	(0.0021)	(0.0046)	(0.0010)	(0.0026)
	$\alpha_2^{(2)}$	0.3550	0.3491	0.3510	0.3526	0.3510	0.3535
	0.35	(0.0050)	(0.0009)	(0.0010)	(0.0026)	(0.0010)	(0.0035)
	$\lambda^{(2)}$	5.9940	5.9780	6.0250	5.9630	6.0000	5.9650
	6	(0.0060)	(0.0022)	(0.0250)	(0.0037)	(0.0000)	(0.0035)
M_4	$\alpha_1^{(1)}$	0.8567	0.8548	0.8213	0.8256	0.8083	0.8127
	0.8	(0.0567)	(0.0548)	(0.0213)	(0.0256)	(0.0083)	(0.0127)
	$\alpha_2^{(1)}$	0.1455	0.1532	0.1177	0.1237	0.1059	0.1020
	0.1	(0.0455)	(0.0532)	(0.0177)	(0.0237)	(0.0059)	(0.0020)
	$\lambda^{(1)}$	6.6410	6.9210	6.8600	6.8240	6.9400	6.6520
	7	(0.3590)	(0.0790)	(0.1400)	(0.1760)	(0.0600)	(0.3480)
	$\alpha_1^{(2)}$	0.6501	0.6548	0.6502	0.6508	0.6507	0.6511
	0.65	(0.0001)	(0.0048)	(0.0002)	(0.0008)	(0.0007)	(0.0011)
	$\alpha_2^{(2)}$	0.2014	0.2044	0.2005	0.2001	0.2000	0.2011
	0.2	(0.0014)	(0.0044)	(0.0005)	(0.0.0001)	(0.0000)	(0.0.0011)
	$\lambda^{(2)}$	3.0020	2.9690	3.0010	3.0070	3.0010	2.9850
	3	(0.0020)	(0.0210)	(0.0001)	(0.0070)	(0.0001)	(0.0015)

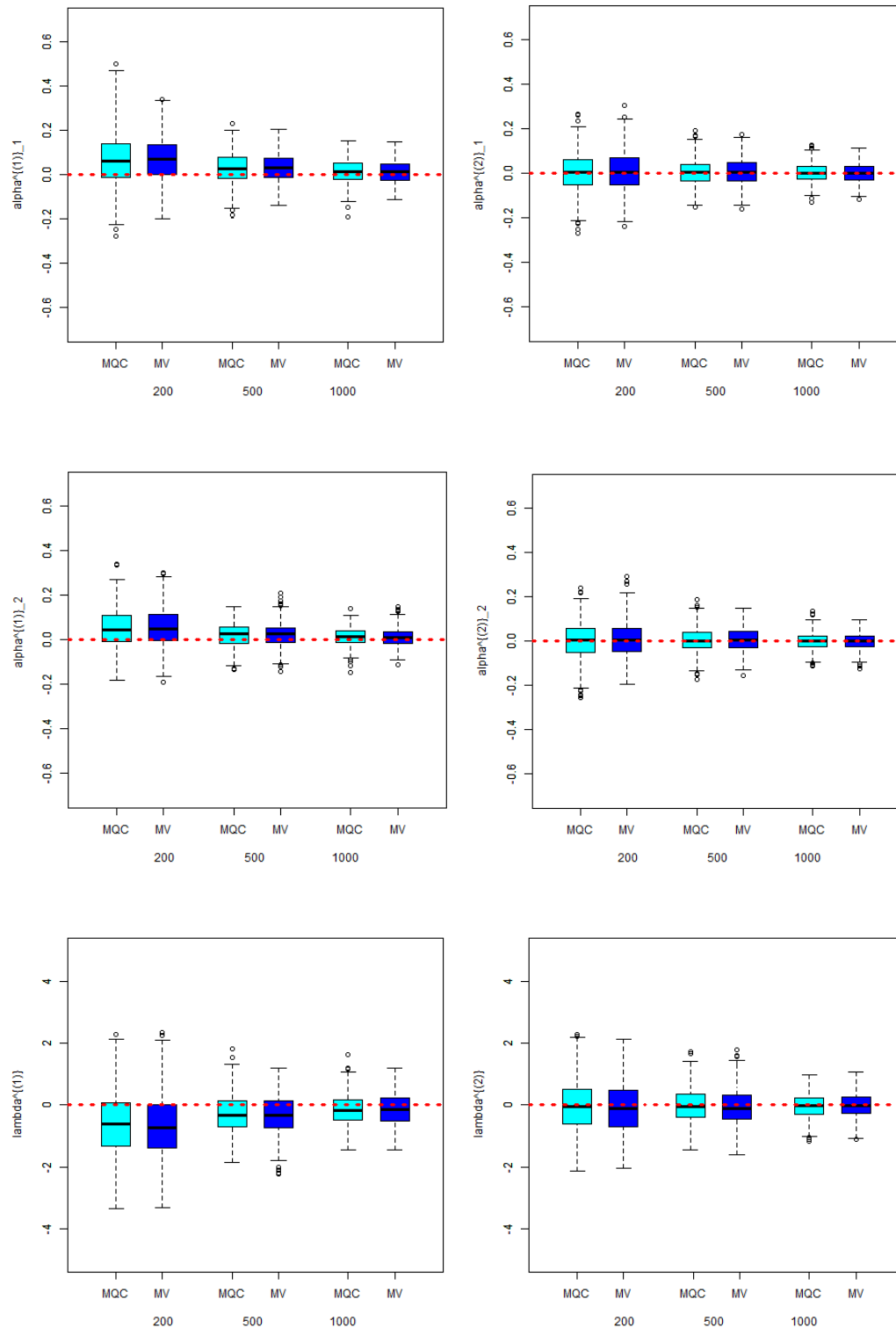


Figura 3.4: Viés das estimativas de $\theta = (0.65, 0.3; 0.6, 0.35; 6; 6)$ do modelo M_3 .

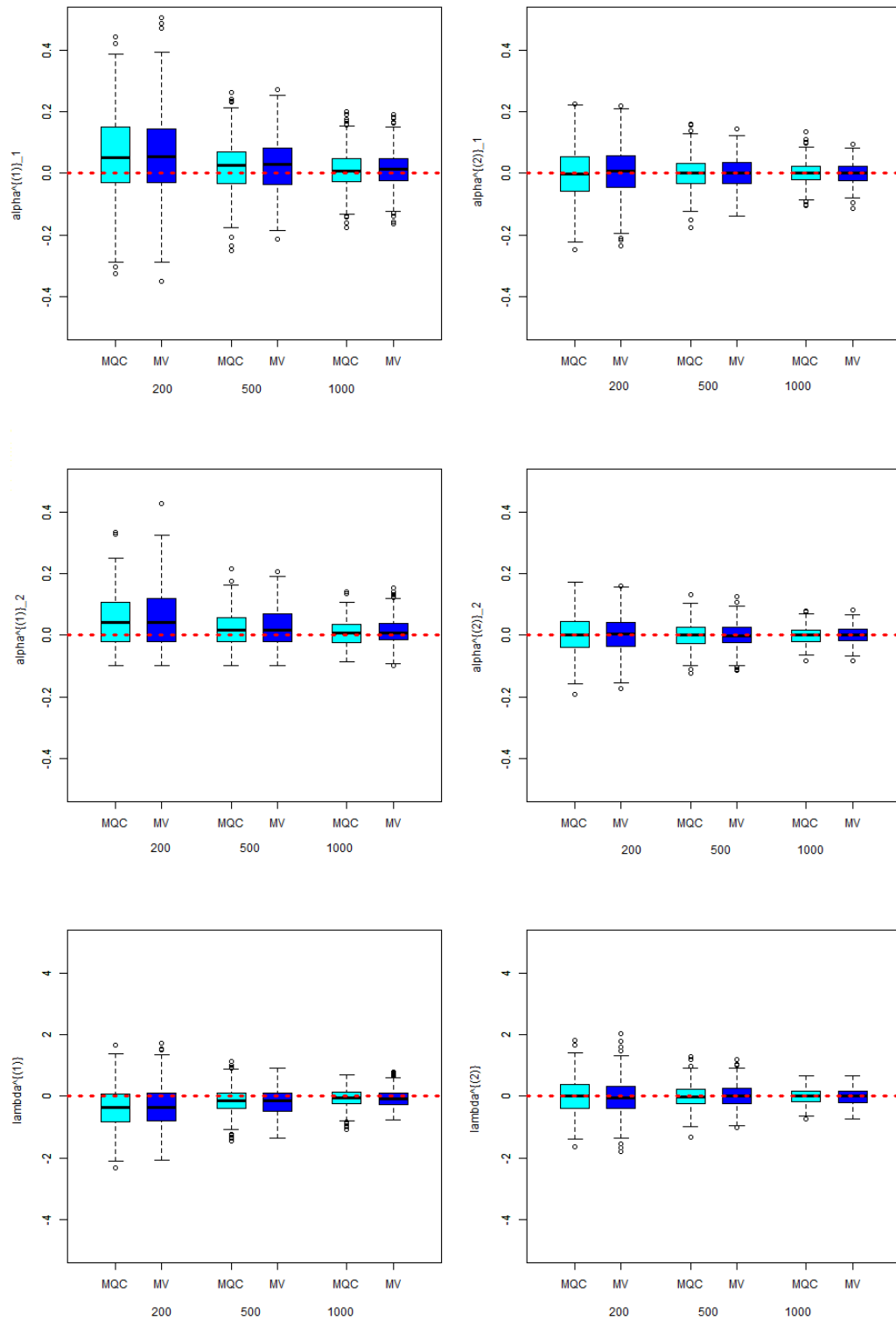


Figura 3.5: Viés das estimativas de $\theta = (0.8, 0.1; 0.65, 0.2; 7; 3)$ do modelo M_4 .

3.4 Predição no modelo PSETINAR(2; 1, 1)_T

Nesta secção pretende-se fazer previsões para observações futuras do modelo definido anteriormente, o modelo PSETINAR(2; 1, 1)_T.

3.4.1 Predição Pontual

Considerando o processo definido em (3.1), pretende-se prever a observação h -passos à frente, X_{i+NT+h} , considerando $h = j + lT$. Com vista à obtenção do estimador pontual ter-se-á de determinar a expressão de X_{i+NT+h} definida à custa das observações atuais X_{i+NT} . Atendendo à definição do modelo e aplicando recursivamente a equação (3.1), tem-se

$$\begin{aligned} X_t &\stackrel{d}{=} \left(\prod_{j=0}^{n-1} \phi_{t-j} \right) \circ X_{t-n} + \sum_{i=1}^{n-1} \left(\prod_{j=0}^{i-1} \phi_{t-j} \right) \circ Z_{t-i} + Z_t \\ &\stackrel{d}{=} \beta_{t,n} \circ X_{t-n} + \sum_{i=0}^{n-1} \beta_{t,i} \circ Z_{t-i} \end{aligned},$$

onde, para $t > i$

$$\beta_{t,i} := \begin{cases} \prod_{j=0}^{i-1} \phi_{t-j}, & \text{se } i > 0 \\ 1, & \text{se } i = 0 \end{cases} = \begin{cases} \beta_{t,j} \beta_{T,T}^l, & \text{se } i = j + lT, j = 1, \dots, T \\ 1, & \text{se } i = 0 \end{cases},$$

pelo que

$$X_{i+NT+h} \stackrel{d}{=} \beta_{i+NT+h,h} \circ X_{i+NT} + \sum_{m=0}^{h-1} \beta_{i+NT+h,m} \circ Z_{i+NT+h-m}.$$

Como $h = j + lT$, tem-se

$$X_{i+NT+h} \stackrel{d}{=} \beta_{i+j+(N+l)T,j+lT} \circ X_{i+NT} + \sum_{m=0}^{j+lT-1} \beta_{i+j+(N+l)T,m} \circ Z_{i+j+(N+l)T-m}.$$

Pela periodicidade dos β 's, $\beta_{i+j+(N+l)T,j+lT} = \beta_{i+j,j+lT} = \beta_{i+j,j}\beta_{T,T}^l$, e considerando a seguinte relação

$$\sum_{m=0}^{lT-1} \beta_{j+lT,m} \circ Z_{j+lT-m} \stackrel{d}{=} \sum_{w=0}^{l-1} \sum_{m=0}^{T-1} \beta_{j+lT,m} \beta_{T,T}^l \circ Z_{j+(l-w)T-m},$$

tem-se, finalmente que

$$X_{i+NT+h} \stackrel{d}{=} (\beta_{i+j,j}\beta_{T,T}^l) \circ X_{i+NT} + V_{i+j+lT},$$

onde

$$V_{i+j+lT} := \sum_{m=0}^{i-1} \beta_{i+j,m} \circ Z_{i+j-m+NT} + \sum_{w=0}^{l-1} \sum_{m=0}^{T-1} \beta_{i+j+(N+l)T,m+j+Tw} \circ Z_{i+(N+l-w)T-m}.$$

Uma vez que o preditor pontual h -passos à frente que minimiza o EQM é dado por

$$\hat{X}_{i+NT+h} = E[X_{i+NT+h}|X_{i+NT}],$$

tem-se

$$\begin{aligned} \hat{X}_{i+NT+h} &= E[(\beta_{i+j,j}\beta_{T,T}^l) \circ X_{i+NT}|X_{i+NT}] \\ &+ \sum_{m=0}^{j-1} \beta_{i+j,m} \left(\lambda_{i+j-m}^{(1)} p_{i+j-m}^{(1)} + \lambda_{i+j-m}^{(2)} p_{i+j-m}^{(2)} \right) \\ &+ \sum_{w=0}^{l-1} \sum_{m=0}^{T-1} \beta_{i+j,m+j+Tw} \left(\lambda_{i-m}^{(1)} p_{i-m}^{(1)} + \lambda_{i-m}^{(2)} p_{i-m}^{(2)} \right), \end{aligned}$$

com

$$\begin{aligned} p_{i+j-m}^{(1)} &= P(X_{i+j-m+NT-1} \leq r_{i+j-m}), \quad p_{i+j-m}^{(2)} = 1 - p_{i+j-m}^{(1)}, \quad i+j > m, \\ p_{i-m}^{(1)} &= P(X_{i+(N+l-w)T-m-1} \leq r_{i-m}), \quad p_{i-m}^{(2)} = 1 - p_{i-m}^{(1)}, \quad i > m. \end{aligned}$$

Suponha-se como caso particular a predição um passo à frente, i.e., $h = 1$. Neste caso,

$$X_{j+NT+1} = \left(\alpha_{j+1}^{(1)} \circ X_{j+NT} + Z_{j+NT+1}^{(1)} \right) I_{j+1}^{(1)} + \left(\alpha_{j+1}^{(2)} \circ X_{j+NT} + Z_{j+NT+1}^{(2)} \right) I_{j+1}^{(2)}.$$

A distribuição preditiva a um passo é dada pela expressão

$$P(X_{j+NT+1} = y | X_{j+NT} = x) = \sum_{m=0}^{\min(x,y)} \sum_{k=1}^2 C_m^x \left(\alpha_{j+1}^{(k)} \right)^m \left(1 - \alpha_{j+1}^{(k)} \right)^{x-m} e^{-\lambda_{j+1}^{(k)}} \frac{\left(\lambda_{j+1}^{(k)} \right)^{(y-m)}}{(y-m)!} I_{j+1}^{(k)},$$

onde $\lambda_{j+1}^{(k)} = E \left[Z_{j+NT+1}^{(k)} \right]$, $k = 1, 2$. O preditor mais usual é o definido pela esperança da distribuição anterior, i.e., dado por

$$\begin{aligned} \hat{X}_{j+NT+1} &= \left(\alpha_{j+1}^{(1)} X_{j+NT} + \lambda_{j+1}^{(1)} \right) P(X_{i+NT} \leq r_{i+1}) \\ &+ \left(\alpha_{j+1}^{(2)} X_{j+NT} + \lambda_{j+1}^{(2)} \right) P(X_{i+NT} > r_{i+1}). \end{aligned}$$

Como preditores ainda se têm a mediana e a moda da distribuição preditiva, que minimizam o erro absoluto médio (condicional) e a função perda tudo ou nada, respetivamente.

3.4.2 Estudo de Simulação

Para analisar e comparar os previsores determinados anteriormente, utilizam-se os modelos apresentados no estudo de simulação da Secção 3.3, especificados na seguinte tabela:

Tabela 3.4: Parâmetros dos modelos $\text{PSETINAR}(2; p^{(1)}, p^{(2)})$

	$\alpha_1^{(1)}$	$\alpha_2^{(1)}$	$\alpha_1^{(2)}$	$\alpha_2^{(2)}$	$\lambda^{(1)}$	$\lambda^{(2)}$	$r^{(1)}$	$r^{(2)}$
M ₁	0.2	0.65	0.2	0.1	7	3	8	4
M ₂	0.2	0.1	0.2	0.1	7	3	8	4
M ₃	0.6	0.3	0.6	0.35	6	6	9	12
M ₄	0.8	0.1	0.65	0.2	7	3	8	4

Recorda-se que o número de réplicas admitidas para cada modelo é 500 e simularam-se 1000 observações para cada réplica. As amostras utilizadas para o estudo de previsão são as n primeiras observações de cada réplica $n = 2N = 200, 400, 800$.

O procedimento geral realizado para a previsão à um passo pode ser descrito da

seguinte forma: a partir de uma amostra de comprimento n , estima-se os valores dos parâmetros a partir dos estimadores de máxima verosimilhança e realiza-se a previsão de $n + 1$; a nova amostra agora é de comprimento $n + 1$, estima-se os novos valores dos parâmetros e prevê-se $n + 2$ e assim sucessivamente. Consideraram-se as estimativas de máxima verosimilhança por serem as estimativas que apresentaram melhores resultados de estimações anteriormente.

Com o objetivo de se comparar os três preditores, pode-se usar a medida dada pelo erro absoluto percentual médio (EAPM) dado por

$$EAPM = \frac{1}{E} \sum_{h=1}^E \frac{|X_{t+1}^* - X_{t+1}|}{X_t},$$

onde X_{t+1}^* representa o preditor de X_{t+1} e E o número de previsões efetuadas.

Para comparar preditores entre diferentes modelos e dimensões utilizaram-se o erro quadrático médio (EQM) para a média condicional, o erro absoluto médio (EAM) para a mediana e a função perda tudo ou nada (FPTN) para a moda dada por

$$FPTN = \frac{1}{m} \sum_{T=1}^m I(x_{t+1}),$$

onde m é o número de réplicas e

$$I(x_{t+1}) = \begin{cases} 1 & \text{se } |\hat{m}o_{t+1} - x_{t+1}| > 1 \\ 0 & \text{se } |\hat{m}o_{t+1} - x_{t+1}| \leq 1 \end{cases}.$$

Na Tabela 3.5 apresentam-se as medidas EQM para a média condicional (\hat{X}_{t+1}), EAM para a mediana (\hat{m}_{t+1}) e FPTN para moda ($\hat{m}o_{t+1}$) a partir de 20 previsões 1-passo à frente para todos os modelos e todas as dimensões consideradas. Percebe-se que para os preditores da média e da moda diminuem à medida que a dimensão da amostra aumenta.

Analisando os valores do erro absoluto percentual médio na Tabela 3.6, pode-se

concluir que, para os modelos analisados, a mediana da distribuição produz melhores resultados que os outros previsores. O modelo M_3 , o único com $\lambda^{(1)} \neq \lambda^{(2)}$, apresenta o EAPM, para as estimativas da média e da mediana, maior quando $N = 200$ do que quando $N = 100$, voltando a diminuir quando $N = 400$.

Tabela 3.5: EQM, EAM e FPTN de 20 predições uma passo à frente para M1, M2 e M3

		\hat{X}_{t+1}	\hat{m}_{t+1}	\hat{mo}_{t+1}
		EQM	EAM	FPTN
M ₁	N=100	9.0616	2.2	0.8
	N=200	8.4229	1.0	0.4
	N=400	7.6240	1.8	0.6
M ₂	N=100	15.8232	2.2	0.8
	N=200	15.0285	2.0	0.6
	N=400	9.6190	1.8	0.4
M ₃	N=100	21.3561	3.4	0.7
	N=200	14.0241	2.9	0.7
	N=400	9.009	2.1	0.6
M ₄	N=100	12.2791	3.3	0.6
	N=200	16.2878	3.4	0.7
	N=400	14.1214	2.1	0.6

Tabela 3.6: EAPM de 20 predições uma passo à frente para M1, M2 e M3

		\hat{X}_{t+1}	\hat{m}_{t+1}	\hat{mo}_{t+1}
M ₁	N=100	0.7738	0.6736	0.6986
	N=200	0.5456	0.3723	0.4390
	N=400	0.3443	0.3041	0.2546
M ₂	N=100	0.7751	0.4502	0.5619
	N=200	0.5851	0.4100	0.4458
	N=400	0.4047	0.4007	0.4142
M ₃	N=100	0.4101	0.5958	0.5355
	N=200	0.4591	0.3239	0.3551
	N=400	0.4047	0.2923	0.2932
M ₄	N=100	0.4595	0.6999	0.6979
	N=200	0.3927	0.3987	0.3104
	N=400	0.2887	0.2825	0.2983

Percebeu-se que quanto maior a dimensão da amostra o erro absoluto percentual médio diminui, mas os valores preditos nem sempre acompanham o comportamento

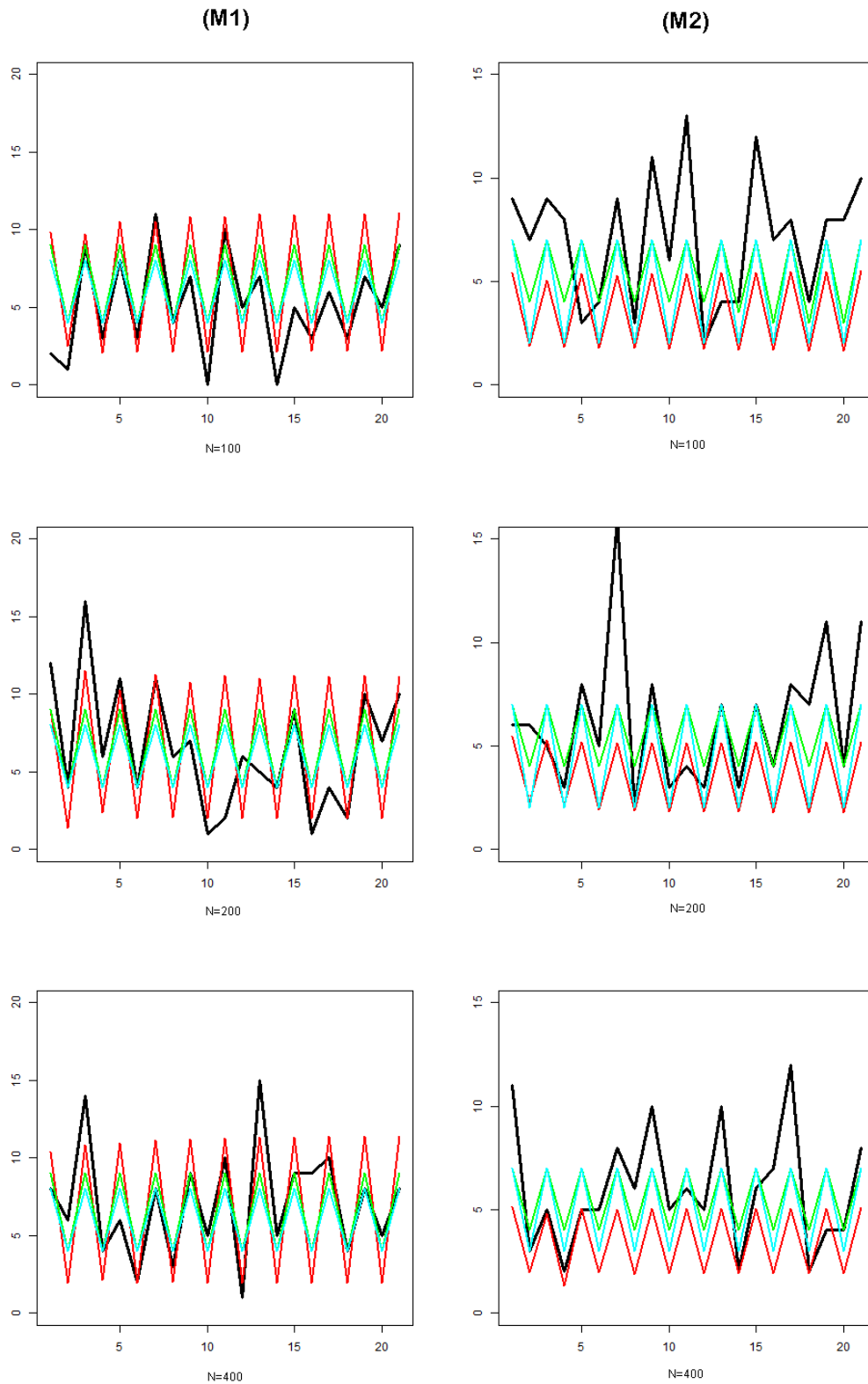


Figura 3.6: Valores de x_{t+1} (preto), \hat{x}_{t+1} (vermelho), \hat{m}_{t+1} (verde) e \hat{m}_{t+1} (azul) de uma realização de M1 e M2.

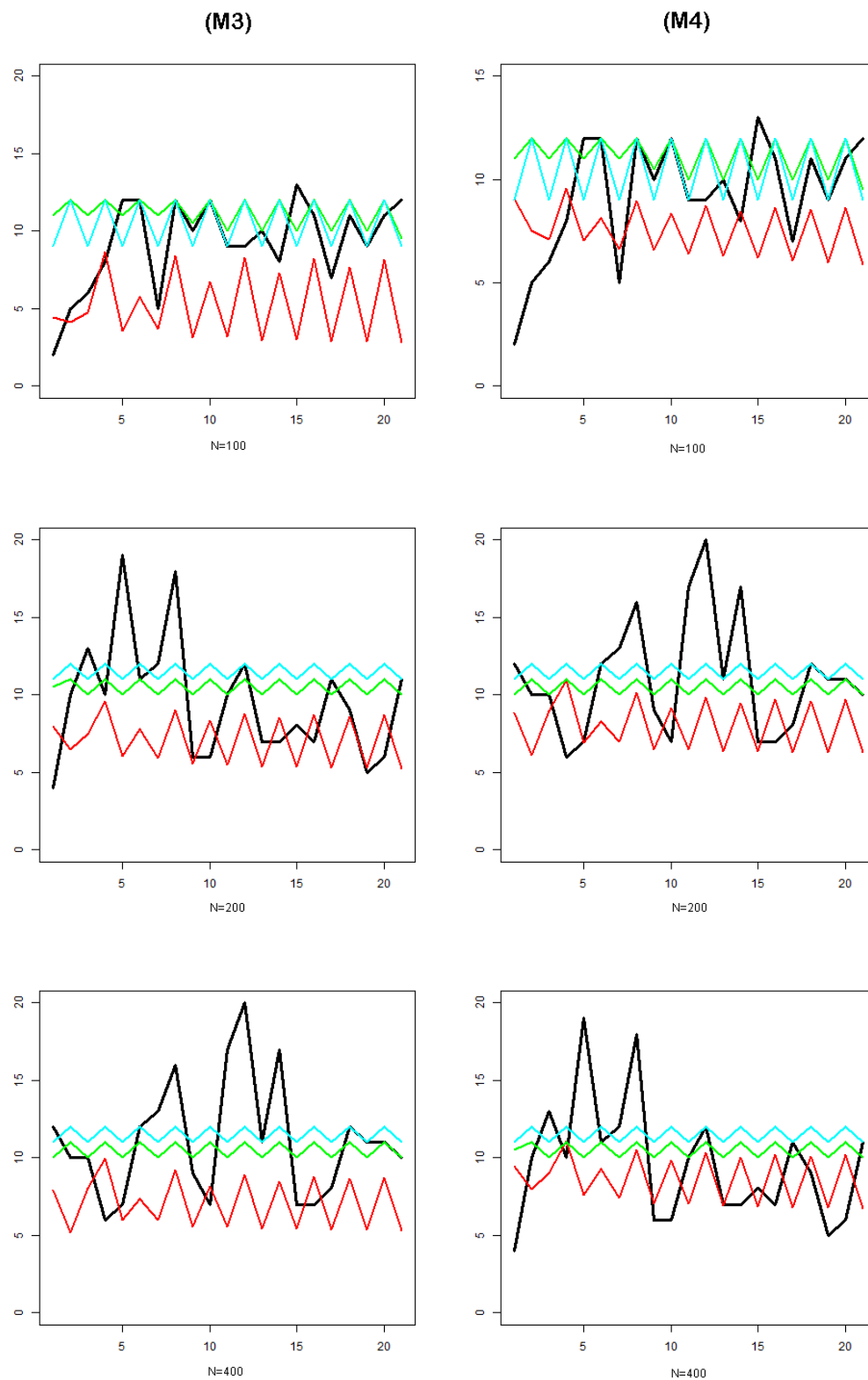


Figura 3.7: Valores de x_{t+1} (preto), \hat{x}_{t+1} (vermelho), \hat{m}_{t+1} (verde) e \hat{m}_{t+1} (azul) de uma realização de M3 e M4.

do real valor dos dados. Na Figura 3.6 pode-se observar o comportamento dos valores preditos para os modelos M_1 e M_2 . Em ambos os modelos as previsões baseadas na média, moda e mediana são bastante próximas. Quando os valores dos parâmetros $\alpha_1^{(k)}$, $k = 1, 2$, são valores maiores que 0.5 como nos modelos M_3 e M_4 , os valores preditos pela moda e mediana são bastante diferentes dos valores preditos baseados na média. Estes modelos tendem a oscilar bastante entre um ramo e outro em cada um dos períodos, o que pode talvez possa justificar o comportamento dos dados preditos apresentados na Figura 3.7.

3.5 Conclusões

Neste capítulo apresentou-se o modelo autorregressivo de valores inteiros de estrutura periódica com limiares autoinduzidos de primeira ordem em cada ramo e período T , designado por $\text{PSETINAR}(2; 1, 1)_T$, cujas inovações são sucessões periódicas de variáveis aleatórias independentes com distribuição de Poisson. Verificou-se que o processo é estritamente estacionário. Foram apresentados os métodos de estimação dos mínimos quadrados condicionais e de máxima verossimilhança condicional e estudadas as suas propriedades assintóticas. Realizou-se um estudo de simulação tendo em vista a comparação de modelos com diferentes valores de parâmetros. Apresentaram-se previsores pontuais, h passos à frente e privilegiou-se a previsão um passo à frente, verificando-se que a mediana amostral da distribuição preditiva produz em média melhores resultados.

Capítulo 4

Considerações Finais

“A mathematician is a blind man in a dark room looking for a black cat which isn’t there.”

Charles Darwin

Esta tese focou-se na análise de modelos de séries de contagem pertencentes à classe de modelos autorregressivos definidos por dois patamares autoinduzidos, com dois regimes, baseados no operador *thinning* binomial e considerando para o processo das inovações sequências de variáveis aleatórias independentes e identicamente distribuídas com distribuição de Poisson.

O Capítulo 1 incidiu na abordagem de alguns conceitos teóricos relevantes para o desenvolvimento deste trabalho, de um modo particular focando diferentes aspetos estudados na literatura por diversos autores sobre os modelos **INAR** e suas extensões.

No Capítulo 2 apresentou-se a classe de modelos autorregressivos de valores inteiros definido por dois limiares autoinduzidos designado por **SETINAR**(2; $p^{(1)}, p^{(2)}$), que apresenta dois regimes cujas ordens dos processos autorregressivos de cada ramo são dadas por $p^{(1)}$ e $p^{(2)}$. Um dos objetivos deste trabalho foi estender o estudo realizado por Monteiro (2010) e Monteiro *et al.* (2012) no qual foi considerado a situação particular de as ordens $p^{(1)}$ e $p^{(2)}$ serem ambas unitárias. Após se terem estabelecido propriedades do modelo e condições para a ergodicidade do processo,

determinaram-se os estimadores de mínimos quadrados condicionais e respectivas propriedades assintóticas e os estimadores de máxima verosimilhança condicional, para os parâmetros. Ainda sob o conhecimento das ordens de cada um dos modelos dos regimes determinaram-se as estimativas dos parâmetros de cada um dos ramos usando a metodologia bayesiana. Verificou-se, através de um estudo de simulação efetuado, que as estimativas bayesianas apresentam um melhor desempenho. Ainda na perspectiva bayesiana, e pretendendo estimar também as ordens dos modelos de cada um dos regimes no caso de elas serem desconhecidas, usou-se o algoritmo RJMCMC, permitindo simultaneamente estimar parâmetros e fazer a seleção de modelos. Como se viu, isto é possível pelo facto de o algoritmo RJMCMC ser uma generalização do algoritmo de Metropolis-Hastings para um espaço de estados mais geral, permitindo a variação da dimensão do espaço do parâmetro entre diferentes iterações da cadeia de Markov usada. Este problema da seleção de modelos e consequente determinação das ordens $p^{(1)}$ e $p^{(2)}$ foi também abordado através de um estudo de simulação, usando os critérios tradicionalmente utilizados, o BIC e o DIC. Os resultados obtidos pelos diferentes métodos utilizados para a seleção de modelos foram sempre concordantes entre si e corroboraram a escolha do modelo correto. No entanto, esta dupla funcionalidade (estimação de parâmetros e seleção de modelos) do algoritmo RJMCMC, que o torna muito atrativo, foi também acompanhada pelo bom desempenho na própria seleção do modelo, como se viu no estudo de simulação feito. Uma outra forma de analisar os resultados obtidos pelo RJMCMC poderia ser através das respectivas distribuições *a posteriori* dos parâmetros $p^{(1)}$ e $p^{(2)}$, em lugar de se calcular as percentagens associadas à escolha de cada modelo. Conclui-se este capítulo com duas aplicações a dados reais considerando séries “históricas” de número de manchas solares e, em ambos os casos, os 3 métodos utilizados para a seleção de modelos (DIC, BIC e RJ) escolheram o mesmo modelo.

No Capítulo 3 propôs-se um modelo autorregressivo de valores inteiros, com estrutura periódica de período T , definido por limiares autoinduzidos de primeira or-

dem em cada ramo, designado por $\text{PSETINAR}(2; 1, 1)_T$, cujas inovações são sucessões também periódicas de variáveis aleatórias independentes com distribuição de Poisson. Apresentaram-se os métodos de estimação dos mínimos quadrados condicionais e de máxima verosimilhança condicional para os parâmetros do modelo em estudo e foram estudadas as suas propriedades assintóticas. No estudo de simulação realizado constatou-se que, apesar de ambos os métodos apresentarem resultados semelhantes, o método de máxima verosimilhança condicional teve um desempenho ligeiramente superior. Esta conclusão é a esperada devido ao facto de se estarem a usar amostras com dimensão elevada, estando em conformidade com as propriedades assintóticas dos métodos. O problema da predição pontual também foi abordado. Deduziu-se a expressão para o preditor h passos à frente que minimiza o EQM e estudou-se com mais detalhe a situação de $h = 1$. No estudo de simulação feito para a predição da observação futura um passo à frente, verificou-se que a mediana da distribuição amostral produziu, de uma forma geral, melhores resultados.

Modelos periódicos com limiares autoinduzidos ainda apresentam inúmeras possibilidades de estudos, um das quais é a ampliação do número de regimes assim como a ordem de cada um deles também pode ser generalizada. Este tópico pode ser sugerido como um futuro trabalho de investigação.

Apêndice **A**

Cálculos Auxiliares

A.1 Algoritmo EM de Monte Carlo

A derivação parcial de Q em relação aos parâmetros $\alpha^{(k)}$ e $\lambda^{(k)}$, com $k = 1, 2$ e $i = \{1, \dots, p^{(k)}\}$, resulta em

$$\begin{cases} \frac{\partial Q}{\partial \alpha_i^{(k)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \frac{y_{t,i}^{(k)} - \alpha_i^{(k)} x_{t-i}}{\alpha_i^{(k)} (1 - \alpha_i^{(k)})} I_{t-d}^{(k)} \\ \frac{\partial Q}{\partial \lambda^{(k)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \left\{ \left[-1 + \frac{z_t^{(k)}}{\lambda^{(k)}} \right] \times I_{t-d}^{(k)} \right\} \end{cases} \quad (\text{A.1})$$

Para verificar (A.1) deriva-se parcialmente Q em relação a cada um dos parâmetros.

Aqui inicia-se a derivação pelo parâmetro $\alpha_1^{(1)}$,

$$\begin{aligned} \frac{\partial Q}{\partial \alpha_1^{(1)}} &= \frac{\partial}{\partial \alpha_1^{(1)}} \left\{ \frac{1}{m} \sum_{j=1}^m l(\theta)_j \right\} = \frac{\partial}{\partial \alpha_1^{(1)}} \left\{ \frac{1}{m} [l(\theta)_1 + l(\theta)_2 + \dots + l(\theta)_m] \right\} \\ &= \frac{1}{m} \left\{ \frac{\partial l(\theta)_1}{\partial \alpha_1^{(1)}} + \frac{\partial l(\theta)_2}{\partial \alpha_1^{(1)}} + \dots + \frac{\partial l(\theta)_m}{\partial \alpha_1^{(1)}} \right\} = \frac{1}{m} \left\{ \sum_{j=1}^m \frac{\partial l(\theta)_j}{\partial \alpha_1^{(1)}} \right\}. \end{aligned}$$

Cada $\frac{\partial l(\theta)_j}{\partial \alpha_1^{(1)}}$ com $1 \leq j \leq m$ é dado por

$$\begin{aligned}
\frac{\partial l(\theta)}{\partial \alpha_1^{(1)}} &= \sum_{t=1}^n \left\{ I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}}{z_t^{(1)}!} \prod_{i=1, i \neq 1}^{p^{(1)}} \left[\binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}} \right] \binom{x_{t-1}}{y_{t,1}^{(1)}} \right. \\
&\quad \left[y_{t,1}^{(1)} \alpha_1^{(1)y_{t,1}^{(1)}-1} (1 - \alpha_1^{(1)})^{x_{t-1}-y_{t,1}^{(1)}} - \alpha_1^{(1)y_{t,1}^{(1)}} (1 - \alpha_1^{(1)})^{x_{t-1}-y_{t,1}^{(1)}-1} (x_{t-1} - y_{t,1}^{(1)}) \right] \times \left\{ \frac{1}{l_1 + l_2} \right\} \Big\} \\
&= \sum_{t=1}^n \left\{ I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}}{z_t^{(1)}!} \prod_{i=1, i \neq 1}^{p^{(1)}} \left[\binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}} \right] \binom{x_{t-1}}{y_{t,1}^{(1)}} \right. \\
&\quad \left. \alpha_1^{(1)y_{t,1}^{(1)}} (1 - \alpha_1^{(1)})^{x_{t-1}-y_{t,1}^{(1)}} \left[\frac{y_{t,1}^{(1)}}{\alpha_1^{(1)}} - \frac{(x_{t-1} - y_{t,1}^{(1)})}{(1 - \alpha_1^{(1)})} \right] \times \left\{ \frac{1}{l_1 + l_2} \right\} \right\}.
\end{aligned} \tag{A.2}$$

Dada a função indicatriz descrita em (2.2) tem-se que para cada $(x_{t-d} \leq r)$ a derivada parcial de Q em relação a $\alpha_1^{(1)}$, apresentada em (A.2), é

$$\begin{aligned}
\frac{\partial l(\theta)}{\partial \alpha_1^{(1)}} &= I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}}{z_t^{(1)}!} \prod_{i=1}^{p^{(1)}} \binom{x_{t-1}}{y_{t,1}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}} \\
&\quad \times \left[\frac{y_{t,1}^{(1)}}{\alpha_1^{(1)}} - \frac{(x_{t-1} - y_{t,1}^{(1)})}{(1 - \alpha_1^{(1)})} \right] \\
&\quad \times \frac{1}{I_{t-d}^{(1)} \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}}{z_t^{(1)}!} \prod_{i=1}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}.
\end{aligned}$$

então

$$\begin{aligned}
 \frac{\partial l(\theta)}{\partial \alpha_1^{(1)}} &= \sum_{t=1}^n \left[\frac{y_{t,1}^{(1)}}{\alpha_1^{(1)}} - \frac{x_{t-1} - y_{t,1}^{(1)}}{(1 - \alpha_1^{(1)})} \right] I_{t-d}^{(1)} \\
 &= \sum_{t=1}^n \frac{y_{t,1}^{(1)} - \alpha_1^{(1)} y_{t,1}^{(1)} - \alpha_1^{(1)} x_{t-1} + \alpha_1^{(1)} y_{t,1}^{(1)}}{\alpha_1^{(1)} (1 - \alpha_1^{(1)})} I_{t-d}^{(1)} \\
 &= \sum_{t=1}^n \frac{y_{t,1}^{(1)} - \alpha_1^{(1)} x_{t-1}}{\alpha_1^{(1)} (1 - \alpha_1^{(1)})} I_{t-d}^{(1)}.
 \end{aligned}$$

Voltando à equação (A.1), tem-se

$$\frac{\partial Q}{\partial \alpha_1^{(1)}} = \frac{1}{m} \sum_{j=1}^m \frac{\partial l(\theta)_j}{\partial \alpha_1^{(1)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \frac{y_{t,1}^{(1)} - \alpha_1^{(1)} x_{t-1}}{\alpha_1^{(1)} (1 - \alpha_1^{(1)})} I_{t-d}^{(1)}. \quad (\text{A.3})$$

A derivação descrita em (A.3) é realizada para todos os $\alpha_i^{(1)}$, com $i = \{1, \dots, p^{(1)}\}$ e desta forma pode-se generalizar como

$$\frac{\partial Q}{\partial \alpha_i^{(1)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \frac{y_{t,i}^{(1)} - \alpha_i^{(1)} x_{t-i}}{\alpha_i^{(1)} (1 - \alpha_i^{(1)})} I_{t-d}^{(1)}. \quad (\text{A.4})$$

Igualando cada uma das derivadas de (A.4) a zero, $\frac{\partial Q}{\partial \alpha_i^{(1)}} = 0$, encontra-se o estimador de máxima verosimilhança de cada um dos parâmetros $\alpha_i^{(1)}$,

$$\hat{\alpha}_i^{(1)} = \frac{\sum_{j=1}^m \sum_{t=1}^n y_{t,i}^{(1)} I_{t-d}^{(1)}}{\sum_{j=1}^m \sum_{t=1}^n x_{t-i} I_{t-d}^{(1)}}. \quad (\text{A.5})$$

Da mesma forma que foram obtidas as equações anteriores para $\alpha_i^{(1)}$, dada a função indicatriz descrita em (2.2) tem-se que para cada $(x_{t-d} > r)$ a derivada parcial de Q em relação a $\alpha_i^{(2)}$ é dada por

$$\frac{\partial Q}{\partial \alpha_i^{(2)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \frac{y_{t,i}^{(2)} - \alpha_i^{(2)} x_{t-i}}{\alpha_i^{(2)} (1 - \alpha_i^{(2)})} I_{t-d}^{(2)}$$

e o estimador de máxima verosimilhança de cada um dos parâmetros $\alpha_i^{(2)}$ é:

$$\hat{\alpha}_i^{(2)} = \frac{\sum_{j=1}^m \sum_{t=1}^n y_{t,i}^{(2)} I_{t-d}^{(2)}}{\sum_{j=1}^m \sum_{t=1}^n x_{t-i} I_{t-d}^{(2)}}. \quad (\text{A.6})$$

Dadas as equações (A.5) e (A.6) a forma geral do o estimador de α é definido por

$$\hat{\alpha}_i^{(k)} = \frac{\sum_{j=1}^m \sum_{t=1}^n y_{t,i}^{(k)} I_{t-d}^{(k)}}{\sum_{j=1}^m \sum_{t=1}^n x_{t-i} I_{t-d}^{(k)}}.$$

Derivando parcialmente Q em relação a $\lambda^{(1)}$ tem-se

$$\begin{aligned} \frac{\partial Q}{\partial \lambda^{(1)}} &= \frac{\partial}{\partial \lambda^{(1)}} \left\{ \frac{1}{m} \sum_{j=1}^m l(\theta)_j \right\} = \frac{\partial}{\partial \lambda_1} \left\{ \frac{1}{m} [l(\theta)_1 + l(\theta)_2 + \cdots + l(\theta)_m] \right\} \\ &= \frac{1}{m} \left\{ \frac{\partial l(\theta)_1}{\partial \lambda^{(1)}} + \frac{\partial l(\theta)_2}{\partial \lambda^{(1)}} + \cdots + \frac{\partial l(\theta)_m}{\partial \lambda^{(1)}} \right\} = \frac{1}{m} \sum_{j=1}^m \frac{\partial l(\theta)_j}{\partial \lambda^{(1)}}. \end{aligned} \quad (\text{A.7})$$

Cada $\frac{\partial l(\theta)_j}{\partial \lambda^{(1)}}$, para $1 \leq j \leq m$, é dado por

$$\begin{aligned} \frac{\partial l(\theta)_j}{\partial \lambda^{(1)}} &= \sum_{t=1}^n I_{t-d}^{(1)} \frac{\prod_{i=1}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}{z_t^{(1)}!} \\ &\quad \times \left(-e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}} + e^{(-\lambda^{(1)})} z_t^{(1)} \lambda^{(1)z_t^{(1)}-1} \right) \times \left\{ \frac{1}{l_1 + l_2} \right\}. \end{aligned}$$

Dada a função indicatriz descrita em (2.2) tem-se que para cada $(x_{t-d} \leq r)$ a derivada parcial de Q em relação ao parâmetro $\lambda^{(1)}$ resume-se em

$$\begin{aligned}
\frac{\partial l(\theta)}{\partial \lambda^{(1)}} &= \frac{\prod_{i=1}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}{z_t^{(1)}!} \\
&\times \left(-e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}} + e^{(-\lambda^{(1)})} z_t^{(1)} \lambda^{(1)z_t^{(1)}-1} \right) \times \frac{1}{l_1} \times I_{t-d}^{(1)} \\
&= \frac{\prod_{i=1}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}}}{z_t^{(1)}!} \times \left(-e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}} + e^{(-\lambda^{(1)})} z_t^{(1)} \lambda^{(1)z_t^{(1)}-1} \right) \\
&\times \frac{1}{\prod_{i=1}^{p^{(1)}} \binom{x_{t-i}}{y_{t,i}^{(1)}} \alpha_i^{(1)y_{t,i}^{(1)}} (1 - \alpha_i^{(1)})^{x_{t-i}-y_{t,i}^{(1)}} \times \left[e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}} \right]} \times I_{t-d}^{(1)} \\
&= \frac{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}} + e^{(-\lambda^{(1)})} z_t^{(1)} \lambda^{(1)z_t^{(1)}-1}}{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}} \times I_{t-d}^{(1)} \\
&= \frac{-e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}}{e^{(-\lambda^{(1)})} \lambda^{(1)z_t^{(1)}}} \times \left(-1 + \frac{z_t^{(1)}}{\lambda^{(1)}} \right) \times I_{t-d}^{(1)} = \left(-1 + \frac{z_t^{(1)}}{\lambda^{(1)}} \right) \times I_{t-d}^{(1)}.
\end{aligned}$$

Retornando a equação (A.7), tem-se

$$\frac{\partial Q}{\partial \lambda^{(1)}} = \frac{1}{m} \sum_{j=1}^m \frac{\partial l(\theta)_j}{\partial \lambda^{(1)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \left(-1 + \frac{z_t^{(1)}}{\lambda^{(1)}} \right) \times I_{t-d}^{(1)}. \quad (\text{A.8})$$

Igualando a derivada de (A.8) a zero encontra-se o estimador de máxima verossimilhança do parâmetro $\lambda^{(1)}$

$$\hat{\lambda}^{(1)} = \frac{\sum_{j=1}^m \sum_{t=1}^n z_t^{(1)} I_{t-d}^{(1)}}{\sum_{j=1}^m \sum_{t=1}^n I_{t-d}^{(1)}}.$$

Seguindo o raciocínio aplicado até a obtenção de $\lambda^{(1)}$ e dada a função indicatriz descrita em (2.2) tem-se que para cada $(x_{t-d} > r)$ a derivada parcial Q em relação a $\lambda^{(2)}$ é dada por

$$\frac{\partial Q}{\partial \lambda^{(2)}} = \frac{1}{m} \sum_{j=1}^m \frac{\partial l(\theta)_j}{\partial \lambda^{(2)}} = \frac{1}{m} \sum_{j=1}^m \sum_{t=1}^n \left(-1 + \frac{z_t^{(2)}}{\lambda^{(2)}} \right) \times I_{t-d}^{(2)}.$$

Assim, generalizando o estimador de λ , para $k = 1, 2$, tem-se

$$\hat{\lambda}^{(k)} = \frac{\sum_{j=1}^m \sum_{t=1}^n z_t^{(k)} I_{t-d}^{(k)}}{\sum_{j=1}^m \sum_{t=1}^n I_{t-d}^{(k)}}.$$

A.1.1 Distribuições Condicionais Completas

A distribuição condicional completa para o parâmetro $\alpha_i^{(k)}$ com $k = \{1, 2\}$, a partir da equação (??) é obtida na sequência.

$$\begin{aligned} \pi \left(\alpha_i^{(k)} | \mathbf{y}, \mathbf{z}, \lambda, \mathbf{x}, \alpha_{i-}^{(1)}, \alpha_{i-}^{(2)}, p^{(1)}, p^{(2)} \right) &\sim \\ &\sim \prod_{t=1}^n \left\{ I_{t-d}^{(1)} \alpha_i^{(1) y_{t,i}^{(1)}} \left(1 - \alpha_i^{(1)} \right)^{x_{t-i} - y_{t,i}^{(1)}} + I_{t-d}^{(2)} \alpha_i^{(2) y_{t,i}^{(2)}} \left(1 - \alpha_i^{(2)} \right)^{x_{t-i} - y_{t,i}^{(2)}} \right\} \\ &\times \underbrace{\lambda^{(1) A^{(1)}} e^{-B^{(1)} \lambda^{(1)}} p^{(1)} \pi(p^{(1)})}_{(a \text{ prioriis para } k=1)} \times \underbrace{\lambda^{(2) A^{(2)}} e^{-B^{(2)} \lambda^{(2)}} p^{(2)} \pi(p^{(2)})}_{(a \text{ prioriis para } k=2)}. \end{aligned}$$

Como o modelo SETINAR considerado apresenta dois regimes, é necessário que se verifique em que ramo estão os parâmetros α_i , desta forma a distribuição de $\alpha_i^{(k)}$

resume-se em aproximadamente o produto das funções

$$\left\{ \begin{array}{ll} \left(\alpha_i^{(1)} \right)^{y_{t,i}^{(1)}} \left(1 - \alpha_i^{(1)} \right)^{x_{t-i} - y_{t,i}^{(1)}}, & t = 1, \dots, n \quad \text{se} \quad x_{t-d} \leq r \\ \left(\alpha_i^{(2)} \right)^{y_{t,i}^{(2)}} \left(1 - \alpha_i^{(2)} \right)^{x_{t-i} - y_{t,i}^{(2)}}, & t = 1, \dots, n \quad \text{se} \quad x_{t-d} > r \end{array} \right.,$$

sendo por fim

$$\prod_{t=1}^n \left\{ \left(\alpha_i^{(k)} \right)^{y_{t,i}^{(k)}} \left(1 - \alpha_i^{(k)} \right)^{x_{t-i} - y_{t,i}^{(k)}} \right\}, \quad t = 1, \dots, n \quad e \quad \left\{ \begin{array}{ll} k = 1 & \text{se} \quad x_{t-d} \leq r \\ k = 2 & \text{se} \quad x_{t-d} > r \end{array} \right.$$

Dado um $t \in \{1, \dots, n\}$ tal que $x_{t-d} \leq r$ define-se $k = 1$, a distribuição condicional completa para um $\alpha_i^{(1)}$ é da forma:

$$\pi \left(\alpha_i^{(1)} | \mathbf{y}, \mathbf{z}, \lambda, \mathbf{x}, \alpha_{i-}^{(1)}, \alpha_{i-}^{(2)}, \mathbf{p}^{(1)}, \mathbf{p}^{(2)} \right) \propto \left(\alpha_i^{(1)} \right)^{\sum_{t=1}^n y_{t,i}^{(1)} I_{t-d}^{(1)}} \left(1 - \alpha_i^{(1)} \right)^{\sum_{t=1}^n (x_{t-i} - y_{t,i}^{(1)}) I_{t-d}^{(1)}}.$$

logo tem-se que

$$\left\{ \begin{array}{l} \pi \left(\alpha_i^{(1)} | \dots \right) \propto \left(\alpha_i^{(1)} \right)^{\left(\sum_{t=1}^n y_{t,i}^{(1)} I_{t-d}^{(1)} + 1 \right) - 1} \left(1 - \alpha_i^{(1)} \right)^{\left(\sum_{t=1}^n (x_{t-i} - y_{t,i}^{(1)}) I_{t-d}^{(1)} + 1 \right) - 1}; \\ \pi \left(\alpha_i^{(1)} | \dots \right) \sim \text{Beta} \left(1 + \sum_{t=1}^n y_{t,i}^{(1)} I_{t-d}^{(1)}, 1 + \sum_{t=1}^n (x_{t-i} - y_{t,i}^{(1)}) I_{t-d}^{(1)} \right). \end{array} \right.$$

Repetindo o raciocínio anterior, dado um $t \in \{1, \dots, n\}$ tal que $x_{t-d} > r$ define-se $k = 2$, a distribuição condicional completa para um $\alpha_i^{(2)}$ é

$$\pi \left(\alpha_i^{(2)} | \dots \right) \sim \text{Beta} \left(1 + \sum_{t=1}^n y_{t,i}^{(2)} I_{t-d}^{(2)}, 1 + \sum_{t=1}^n (x_{t-i} - y_{t,i}^{(2)}) I_{t-d}^{(2)} \right).$$

De forma geral a distribuição de $\alpha_i^{(k)}$, $k = \{1, 2\}$ e $i = 1, \dots, p^{(k)}$ é dada por:

$$\pi \left(\alpha_i^{(k)} | \mathbf{y}, \mathbf{z}, \lambda, \mathbf{x}, \alpha_{i-}^{(1)}, \alpha_{i-}^{(2)}, p^{(1)}, p^{(2)} \right) \sim \text{Beta} \left(1 + \sum_{t=1}^n y_{t,i}^{(k)} I_{t-d}^{(k)}, 1 + \sum_{t=1}^n (x_{t-i} - y_{t,i}^{(k)}) I_{t-d}^{(k)} \right).$$

A distribuição condicional completa para o parâmetro λ a partir da equação (??) é obtida na sequência.

$$\begin{aligned} \pi \left(\lambda^{(1)} | \mathbf{y}, \mathbf{z}, \mathbf{x}, \alpha_i^{(1)}, \alpha_i^{(2)}, p^{(1)}, p^{(2)} \right) &\propto \\ &\propto \prod_{t=1}^n \left\{ I_{t-d}^{(1)} \alpha_i^{(1) y_{t,i}^{(1)}} \left(1 - \alpha_i^{(1)} \right)^{x_{t-i} - y_{t,i}^{(1)}} + I_{t-d}^{(2)} \alpha_i^{(2) y_{t,i}^{(2)}} \left(1 - \alpha_i^{(2)} \right)^{x_{t-i} - y_{t,i}^{(2)}} \right\} \\ &\times \underbrace{\lambda^{(1) A^{(1)}} e^{-B^{(1)} \lambda^{(1)}} p^{(1)} \pi(p^{(1)})}_{(a \text{ priori para } k=1)} \times \underbrace{\lambda^{(2) A^{(2)}} e^{-B^{(2)} \lambda^{(2)}} p^{(2)} \pi(p^{(2)})}_{(a \text{ priori para } k=2)}. \end{aligned}$$

Como os parâmetros $\lambda^{(k)}$ são independentes, para $t \in \{1, \dots, n\}$ tal que $x_{t-d} \leq r$, define-se $k = 1$, logo a distribuição condicional completa para $\lambda^{(1)}$ é dada por:

$$\begin{aligned} \pi \left(\lambda_t^{(1)} | \dots \right) &\propto \left(\lambda^{(1)} \right)^{A^{(1)}} e^{-B^{(1)} \lambda^{(1)}} \prod_{t=1}^n I_{t-d}^{(1)} e^{(-\lambda^{(1)})} \left(\lambda^{(1)} \right)^{z_t^{(1)}}; \\ \pi \left(\lambda_t^{(1)} | \dots \right) &\sim \left(\lambda^{(1)} \right)^{A^{(1)}} e^{-B^{(1)} \lambda^{(1)}} e^{\left(-\lambda^{(1)} \sum_{t=1}^n I_{t-d}^{(1)} \right)} \left(\lambda^{(1)} \right)^{\left(\sum_{t=1}^n I_{t-d}^{(1)} z_t^{(1)} \right)}; \\ \pi \left(\lambda_t^{(1)} | \dots \right) &\sim \left(\lambda^{(1)} \right)^{A^{(1)} + \left(\sum_{t=1}^n I_{t-d}^{(1)} z_t^{(1)} \right)} e^{\left(\lambda^{(1)} \left(-B^{(1)} + \sum_{t=1}^n I_{t-d}^{(1)} \right) \right)}; \\ \pi \left(\lambda_t^{(1)} | \dots \right) &\propto \left(\lambda^{(1)} \right)^{1 + A^{(1)} + \left(\sum_{t=1}^n I_{t-d}^{(1)} z_t^{(1)} \right)} e^{\left(\lambda^{(1)} \left(-B^{(1)} + \sum_{t=1}^n I_{t-d}^{(1)} \right) \right)}; \\ \pi \left(\lambda_t^{(1)} | \dots \right) &\sim \text{Gama} \left(1 + A^{(1)} + \sum_{t=1}^n I_{t-d}^{(1)} z_t^{(1)}, B^{(1)} + \sum_{t=1}^n I_{t-d}^{(1)} \right). \end{aligned}$$

Repetindo o raciocínio anterior, dado um $t \in \{1, \dots, n\}$ tal que $x_{t-d} > r$ define-se $k = 2$, a distribuição condicional completa para cada $\lambda_t^{(2)}$ é

$$\pi \left(\lambda_t^{(2)} | \dots \right) \sim \text{Gama} \left(1 + A^{(1)} + \sum_{t=1}^n I_{t-d}^{(2)} z_t^{(2)}, B^{(2)} + \sum_{t=1}^n I_{t-d}^{(2)} \right).$$

De uma forma geral a distribuição de $\lambda_t^{(k)}$, $k = \{1, 2\}$ e $i = 1, \dots, p^{(k)}$ é dada por

$$\pi \left(\lambda_t^{(k)} | \mathbf{y}, \mathbf{z}, \mathbf{x}, \alpha_i^{(1)}, \alpha_i^{(2)}, p^{(1)}, p^{(2)} \right) \sim \text{Gama} \left(1 + A^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} z_t^{(k)}, B^{(k)} + \sum_{t=1}^n I_{t-d}^{(k)} \right).$$

A.2 Cálculos para o Modelo PSETINAR

A.2.1 Derivadas Parciais de $\frac{p_j(b-1|a)}{p_j(b|a)}$

$$\begin{aligned}
\frac{\partial}{\partial \alpha_j^{(k)}} \left(\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right) &= \frac{\frac{\partial p_j(b-1|a)^{(k)}}{\partial \alpha_j^{(k)}} p_j(b|a)^{(k)} - \frac{\partial p_j(b|a)^{(k)}}{\partial \alpha_j^{(k)}} p_j(b-1|a)}{p_j(b|a)^{(k)^2}} \\
&= \frac{\frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left[(b-1-\alpha_j^{(k)}a) p_j(b-1|a)^{(k)} - \lambda_j p_j(b-2|a)^{(k)} \right] p_j(b|a)^{(k)}}{p_j(b|a)^{(k)^2}} \\
&= \frac{\frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left[(b-\alpha_j^{(k)}a) p_j(b|a)^{(k)} - \lambda_j p_j(b-1|a)^{(k)} \right] p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)^2}} \\
&= \frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ (a-1-\alpha_j^{(k)}b) \frac{p_j(b-1|a)^{(k)} p_j(b|a)^{(k)}}{p_j(b|a)^{(k)^2}} - \lambda_j \frac{p(b-2|a)^{(k)} p_j(b|a)^{(k)}}{p_j(b|a)^{(k)^2}} \right. \\
&\quad \left. - (b-\alpha_j^{(k)}a) \frac{p(b-1|a)^{(k)} p_j(b|a)^{(k)}}{p_j(b|a)^{(k)^2}} + \lambda_j \frac{p(b-1|a)^{(k)^2}}{p_j(b|a)^{(k)^2}} \right\} \\
&= \frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ (b-1-\alpha_j^{(k)}a) \frac{p(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} - \lambda_j \frac{p(b-2|a)^{(k)}}{p_j(b|a)^{(k)}} \right. \\
&\quad \left. - (b-\alpha_j^{(k)}a) \frac{p(b-1|a)}{p_j(b|a)} + \lambda_j \frac{p(b-1|a)^2}{p_j(b|a)^2} \right\} \\
&= -\frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ \frac{p(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} + \lambda_j \frac{p(b-2|a)^{(k)}}{p_j(b|a)^{(k)}} - \frac{p(b-1|a)^{(k)^2}}{p_j(b|a)^{(k)^2}} \right\}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \lambda_j} \left(\frac{p_j(b-1|a)}{p_j(b|a)} \right) &= \frac{\frac{\partial p_j(b-1|a)}{\partial \lambda_j} p_j(b|a) - \frac{\partial p_j(b|a)}{\partial \lambda_j} p_j(b-1|a)}{p_j(b|a)^2} \\
&= \frac{[p_j(b-2|a) - p_j(b-1|a)] p_j(b|a) - [p_j(b-1|a) - p_j(b|a)] p_j(b-1|a)}{p_j(b|a)^2} \\
&= \frac{p_j(b-2|a) p_j(b|a)}{p_j(b|a)^2} - \frac{p_j(b-1|a) p_j(b|a)}{p_j(b|a)^2} - \frac{p_j(b-1|a)^2}{p_j(b|a)^2} + \frac{p_j(b-1|a) p_j(b|a)}{p_j(b|a)^2} \\
&= \frac{p_j(b-2|a) p_j(b|a)}{p_j(b|a)^2} - \frac{p_j(b-1|a)^2}{p_j(b|a)^2}
\end{aligned}$$

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \alpha_j^{(k)}} \left(\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right) = \frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ \frac{p(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} + \lambda_j \frac{p(b-2|a)^{(k)}}{p_j(b|a)^{(k)}} - \frac{p(b-1|a)^{(k)^2}}{p_j(b|a)^{(k)^2}} \right\} \\ \frac{\partial}{\partial \lambda_j} \left(\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)} \right) = \frac{p_j(b-2|a)}{p_j(b|a)^2} - \frac{p_j(b-1|a)^2}{p_j(b|a)^2} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \alpha_j^{(k)}} \left(\frac{p_j(x_{j+sT}-1|x_{j-1+sT})}{p_j(x_{j+sT}|x_{j-1+sT})} \right) = \frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ \frac{p(x_{j+sT}-1|x_{j-1+sT})}{p_j(x_{j+sT}|x_{j-1+sT})} + \lambda_j \frac{p(x_{j+sT}-2|x_{j-1+sT})}{p_j(x_{j+sT}|x_{j-1+sT})} - \frac{p(x_{j+sT}-1|x_{j-1+sT})^2}{p_j(x_{j+sT}|x_{j-1+sT})^2} \right\} \\ \frac{\partial}{\partial \lambda_j} \left(\frac{p_j(x_{j+sT}-1|x_{j-1+sT})}{p_j(x_{j+sT}|x_{j-1+sT})} \right) = \frac{p_j(x_{j+sT}-2|x_{j-1+sT})}{p_j(x_{j+sT}|x_{j-1+sT})^2} - \frac{p_j(x_{j+sT}-1|x_{j-1+sT})^2}{p_j(x_{j+sT}|x_{j-1+sT})^2} \end{array} \right. .$$

Substituindo $p(\cdot|\cdot)$ pelos seus valores originais, tem-se

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \alpha_j^{(k)}} \left(\frac{p_j \left(x_{j-1+sT}, x_{j+sT}-1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)} \right) = \\ = \frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ \frac{p_j \left(x_{j-1+sT}, x_{j+sT}-1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)} + \lambda_j \frac{p_j \left(x_{j-1+sT}, x_{j+sT}-2, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)} - \frac{p_j \left(x_{j-1+sT}, x_{j+sT}-1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)^2}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)^2} \right\} \\ \frac{\partial}{\partial \lambda_j} \left(\frac{p_j \left(x_{j-1+sT}, x_{j+sT}-1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)} \right) = \\ = \frac{p_j \left(x_{j-1+sT}, x_{j+sT}-2, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)^2} - \frac{p_j \left(x_{j-1+sT}, x_{j+sT}-1, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)^2}{p_j \left(x_{j-1+sT}, x_{j+sT}, \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}, \lambda_j \right)^2} \end{array} \right. .$$

A.2.2 Derivadas Parciais de Segunda Ordem do Logaritmo da Função de Verossimilhança

$$\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)2}} &= \frac{\partial}{\partial \alpha_j^{(k)}} \left(\frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_{s=0}^{N-1} (b - a\alpha_j^{(k)}) - \lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} I_j^{(k)} \right) \\
&= -\frac{1 - 2\alpha_j^{(k)}}{\alpha_j^{(k)2}(1 - \alpha_j^{(k)})^2} \sum_{s=1}^n \left\{ (b - a\alpha_j^{(k)}) - \lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right\} I_j^{(k)} + \frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \\
&\quad \sum_{s=1}^n \left\{ -a + \frac{\lambda_j}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \left[\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} + \lambda_j \frac{p_j(b-2|a)^{(k)}}{p_j(b|a)^{(k)}} - \lambda_j \frac{p_j(b-1|a)^{(k)2}}{p_j(b|a)^{(k)2}} \right] \right\} \\
&= -\frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_{s=1}^n \left\{ -b + a\alpha_j^{(k)} + 2b\alpha_j^{(k)} - 2a\alpha_j^{(k)2} + \lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} - 2\alpha_j^{(k)}\lambda_j \right. \\
&\quad \left. \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} - a\alpha_j^{(k)} + a\alpha_j^{(k)2} + \lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} + \lambda_j^2 \frac{p(b-2|a)^{(k)}}{p(b|a)^{(k)}} - \lambda_j^2 \frac{p(b-1|a)^{(k)2}}{p(b|a)^{(k)2}} \right\} \\
&= -\frac{I_j^{(k)}}{\alpha_j^{(k)}(1 - \alpha_j^{(k)})} \sum_{s=1}^n \left\{ (2\alpha_j^{(k)} - 1)b - a\alpha_j^{(k)2} + (1 - 2\alpha_j^{(k)})\lambda_j \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right. \\
&\quad \left. + \lambda_j^2 \frac{p(b-2|a)^{(k)}}{p(b|a)^{(k)}} - \lambda_j^2 \frac{p(b-1|a)^2}{p(b|a)^2} \right\}.
\end{aligned} \tag{A.9}$$

$$\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial \lambda_j^2} &= \frac{\partial}{\partial \lambda_j} \sum_{s=0}^{N-1} \left(\frac{p_j(b-1|a)^{(1)}}{p_j(b|a)^{(1)}} I_j^{(1)} + \frac{p_j(b-1|a)^{(2)}}{p_j(b|a)^{(2)}} I_j^{(2)} - (N-1) \right) \\
&= \sum_{s=1}^n \frac{\partial}{\partial \lambda_j} \left(\frac{p_j(b-1|a)^{(1)}}{p_j(b|a)^{(1)}} I_j^{(1)} + \frac{p_j(b-1|a)^{(2)}}{p_j(b|a)^{(2)}} I_j^{(2)} \right) \\
&= \sum_{s=1}^n \sum_{k=1}^2 \left[\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right] I_j^{(k)}.
\end{aligned} \tag{A.10}$$

$$\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(1)} \partial \alpha_j^{(2)}} = 0, k = 1, 2; \quad j = 1, \dots, T. \tag{A.11}$$

$$\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)} \partial \lambda_j} &= \frac{\partial}{\partial \alpha_j^{(k)}} \sum_{s=0}^{N-1} \left(\frac{p_j(b-1|a)^{(1)}}{p_j(b|a)^{(1)}} I_j^{(1)} + \frac{p_j(b-1|a)^{(2)}}{p_j(b|a)^{(2)}} I_j^{(2)} - (N-1) \right) \\
&= \sum_{s=1}^n \frac{\partial}{\partial \alpha_j^{(k)}} \left(\frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right) I_j^{(k)} \\
&= \sum_{s=1}^n - \frac{1}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ \frac{p_j(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} + \lambda_j \frac{p_j(b-2|a)^{(k)}}{p_j(b|a)^{(k)}} - \frac{p_j(b-1|a)^{(k)^2}}{p_j(b|a)^{(k)^2}} \right\} I_j^{(k)}.
\end{aligned}$$

Para a determinação das estimativas da matriz de Fisher usa-se o método sugerido por Silva *et al.* (2005) que baseia-se nas propriedades do valor médio. Verifica-se nas expressões das derivadas de segunda ordem do logaritmo da função de verosimilhança, equações (A.9), (A.10) e (A.2.2), que para cada j a variar de $1 - T$, as derivadas só dependem de x_{j+sT} e x_{j-1+sT} e portanto calcular a estimativa da matriz de Fischer basta calcular a esperança conjunta usando as estimativas de máxima verosimilhança dos parâmetros no lugar dos verdadeiros valores.

$$\left\{ \begin{array}{l} \frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)2}} = \sum_{s=0}^{N-1} G_{1,j}^{(k)}(X_{j-1+sT}, X_{j+sT}) I_j^{(k)} \\ \frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)} \partial \lambda_j} = \sum_{s=0}^{N-1} G_{2,j}^{(k)}(X_{j-1+sT}, X_{j+sT}) I_j^{(k)} \\ \frac{\partial^2 l(\theta)}{\partial \lambda_j^2} = \sum_{s=0}^{N-1} G_{3,j}^{(k)}(X_{j-1+sT}, X_{j+sT}) I_j^{(k)}. \end{array} \right. \quad (\text{A.12})$$

Assim,

$$\begin{aligned}
E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)2}} \right] &= N \sum_a^{+\infty} \sum_b^{+\infty} P(X_{j+sT} = b) \frac{I_j^{(k)}}{\alpha_j^{(k)}(1-\alpha_j^{(k)})} \left\{ \left[(2\alpha_j^{(k)} - 1)b - a\alpha_j^{(k)^2} \right] \right. \\
&\quad \left. p_j(b|a) + 2(1-\alpha_j^{(k)})\lambda_j p_j(b-1|a)^{(k)} + \lambda_j^2 p_j(b-2|a)^{(k)} - \lambda_j^2 \frac{p_j^2(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right\};
\end{aligned}$$

$$\begin{aligned}
E \left[\frac{\partial^2 l(\theta)}{\partial \alpha_j^{(k)} \partial \lambda_j} \right] &= \frac{N}{\alpha_j^{(k)} (1 - \alpha_j^{(k)})} \sum_a^{+\infty} \sum_b^{+\infty} P(X_{j+sT} = b) \times \\
&\quad \times \left(-p_j(b-1|a)^{(k)} - \lambda_j p_j(b-2|a)^{(k)} + \lambda_j \frac{p_j^2(b-1|a)^{(k)}}{p(b|a)^{(k)}} \right); \\
E \left[\frac{\partial^2 l(\theta)}{\partial} \right] &= N \sum_a^{+\infty} \sum_b^{+\infty} P(X_{j+sT} = b) \left\{ \lambda_j p_j(b-2|a)^{(k)} - \frac{p_j^2(b-1|a)^{(k)}}{p_j(b|a)^{(k)}} \right\}.
\end{aligned}$$

Apêndice B

Funções Implementadas no Programa R

B.1 Funções Utilizadas para o Modelo SETINAR

```
##### Simular um conjuntos de dados SETINAR(2;p,q)#####
simula_setinar=function(alpha1,alpha2,lambda1,lambda2,R,d,burn,n)
{p1<-length(alpha1)
 p2<-length(alpha2)
 pmax=max(p1,p2)
 x<-rep(0,burn+n+d+pmax)
 z=matrix(0,ncol=pmax,nrow=(burn+n+d+pmax))
 phi<-rep(0,burn+n+d+pmax)
 r<-R
 s<-max(p1,p2,d)
 for(i in 1:s)
 { x[i]<-as.integer(mean(c(1/(1-sum(alpha1)),1/(1-sum(alpha2))))))}

 for(i in (s+1+d):(burn+n+d+pmax))
 { y<-0
  log<-(x[i-d]>r)
  if(log=="TRUE")
  {for(j in 1:p2)
   { y[j]=rbinom(1,x[i-j],alpha2[j])+rpois(1,lambda2)}}}
  else
  {for(j in 1:p1)
   { y[j]=rbinom(1,x[i-j],alpha1[j])+rpois(1,lambda1)}}}
  x[i]=sum(y)
  phi[i] }
 resp=x[(burn+1):(burn+n+d+pmax)]
 resp}
#####
```

```
##### Método dos mínimos quadrados condicionais para R fixo
mqc_setinar<-function(x,p1,p2,r,d)
{n<-length(x);
 Ind_1<-rep(0,n-d);
 Ind_2<-rep(0,n-d);
 for(i in (d+1):(n))
 {if(x[i-d]>r){Ind_2[i-d]<-1}
  else{Ind_1[i-d]<-1}}
 prop<-sum(Ind_1)/(n-d)
 if(p1==1 && p2==1)
 { ord<-p1+p2+2;
  Mat_sist<-array(0,dim=c(ord,ord));
  Mat_sist[1,1]<-sum(Ind_1*x[1:(n-1)]^2);
  Mat_sist[2,2]<-sum(Ind_2*x[1:(n-1)]^2);
  Mat_sist[3,3]<-sum(Ind_1);
  Mat_sist[4,4]<-sum(Ind_2);
  Mat_sist[1,3]<-sum(Ind_1*x[1:(n-1)]);
  Mat_sist[2,4]<-sum(Ind_2*x[1:(n-1)]);
  Mat_sist[3,1]<-sum(Ind_1*x[1:(n-1)]);
  Mat_sist[4,2]<-sum(Ind_2*x[1:(n-1)]);
  L<-array(0,dim=c(ord,1));
  L[1,1]<-sum(Ind_1*x[1:(n-1)]*x[2:n]);
  L[2,1]<-sum(Ind_2*x[1:(n-1)]*x[2:n]);
  L[3,1]<-sum(Ind_1*x[2:n]);
  L[4,1]<-sum(Ind_2*x[2:n]);}

 if(p1==1 && p2==2)
 { ord<-p1+p2+2;
  Mat_sist<-array(0,dim=c(ord,ord));
  Mat_sist[1,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]^2);
  Mat_sist[2,2]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]^2);
  Mat_sist[3,3]<-sum(Ind_2[2:(n-1)]*x[1:(n-2)]^2);
  Mat_sist[4,4]<-sum(Ind_1[2:(n-1)]);
  Mat_sist[5,5]<-sum(Ind_2[2:(n-1)]);
  Mat_sist[1,4]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[2,5]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[3,5]<-sum(Ind_2[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[2,3]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[3,2]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[4,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[5,2]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[5,3]<-sum(Ind_2[2:(n-1)]*x[1:(n-2)]);
  L<-array(0,dim=c(ord,1));
  L[1,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]*x[3:n]);
  L[2,1]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]*x[3:n]);
  L[3,1]<-sum(Ind_2[2:(n-1)]*x[1:(n-2)]*x[3:n]);
  L[4,1]<-sum(Ind_1[2:(n-1)]*x[3:n]);
  L[5,1]<-sum(Ind_2[2:(n-1)]*x[3:n]); }
```

```

if(p1==2 && p2==1)
{ ord<-p1+p2+2;
  Mat_sist<-array(0,dim=c(ord,ord));
  Mat_sist[1,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]^2);
  Mat_sist[2,2]<-sum(Ind_1[2:(n-1)]*x[1:(n-2)]^2);
  Mat_sist[3,3]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]^2);
  Mat_sist[4,4]<-sum(Ind_1[2:(n-1)]);
  Mat_sist[5,5]<-sum(Ind_2[2:(n-1)]);
  Mat_sist[1,4]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[2,4]<-sum(Ind_1[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[3,5]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[1,2]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[2,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[4,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]);
  Mat_sist[4,2]<-sum(Ind_1[2:(n-1)]*x[1:(n-2)]);
  Mat_sist[5,3]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]);
  L<-array(0,dim=c(ord,1));
  L[1,1]<-sum(Ind_1[2:(n-1)]*x[2:(n-1)]*x[3:n]);
  L[2,1]<-sum(Ind_1[2:(n-1)]*x[1:(n-2)]*x[3:n]);
  L[3,1]<-sum(Ind_2[2:(n-1)]*x[2:(n-1)]*x[3:n]);
  L[4,1]<-sum(Ind_1[2:(n-1)]*x[3:n]);
  L[5,1]<-sum(Ind_2[2:(n-1)]*x[3:n]);}
if(det(Mat_sist)==0){ p<-rep(0,ord)}
else{ p<-solve(Mat_sist,L); }
if(any(p[1:ord]<0)){p[1:ord]<-NA}
else{if(any(p[1:(p1+p2)]>1)){p[1:ord]<-NA} }
if(p1==1 && p2==1)
{ sqe<-sum((x[2:n]-p[1]*Ind_1*x[1:(n-1)]-p[2]*Ind_2*x[1:(n-1)]
  -p[3]*Ind_1-p[4]*Ind_2)^2)}
if(p1==1 && p2==2)
{ sqe<-sum((x[3:n]-p[1]*Ind_1[2:(n-1)]*x[2:(n-1)]-p[2]*Ind_2[2:(n-1)]
  *x[2:(n-1)]-p[3]*Ind_2[2:(n-1)]*x[1:(n-2)]-p[4]*Ind_1[2:(n-1)]
  -p[5]*Ind_2[2:(n-1)]^2)}
if(p1==2 && p2==1)
{ sqe<-sum((x[3:n]-p[1]*Ind_1[2:(n-1)]*x[2:(n-1)]-p[2]*Ind_1[2:(n-1)]
  *x[1:(n-2)]-p[3]*Ind_2[2:(n-1)]*x[2:(n-1)]-p[4]*Ind_1[2:(n-1)]
  -p[5]*Ind_2[2:(n-1)]^2)}
return(rbind(p,sqe,prop))}
#####

```

Monte Carlo Expectation-Maximization para R fixo

```
em_setinar<-function(x,p1,p2,r,d,run)
```

```
{ n<-length(x);
```

```
  Ind_1<-rep(0,n-d);
```

```
  Ind_2<-rep(0,n-d);
```

```
for(i in (d+1):(n))
```

```
{if(x[i-d]>r){Ind_2[i-d]<-1}
```

```
  else{Ind_1[i-d]<-1} }
```

```
prop<-sum(Ind_1)/(n-d)
```

```
pmax<-max(p1,p2)
```

```
n=length(x)-pmax
```

```

y=matrix(0,ncol=pmax,nrow=n)
y_sum=matrix(0,ncol=1,nrow=n)
y_1m=matrix(0,ncol=100,nrow=n)
y_2m=matrix(0,ncol=100,nrow=n)
y_3m=matrix(0,ncol=100,nrow=n)
x_m=matrix(0,ncol=100,nrow=n)
z=matrix(0,ncol=1,nrow=n)
z_m=matrix(0,ncol=100,nrow=n)
resp=matrix(0,ncol=(p1+p2+2),nrow=run)

# Prioris
alpha1=0; a1.beta=b1.beta=1;
for(i in 1:p1)
{alpha1[i]=0.1 }
alpha2=0; a2.beta=b2.beta=1;
for(i in 1:p2)
{alpha2[i]=0.1 }
a.lambda=1; b.lambda=1
lambda1=rgamma(1,a.lambda,b.lambda)

cont<-1
while(cont<10)
{ somay1=somay21=somay22=somax1=somax21=somax22=matrix(0,ncol=1,nrow=100)
  somaz1=somaz2=somaum1=somaum2=matrix(0,ncol=1,nrow=100)
  m<-1
  while (m<100)
  { for(t in (d+1):n)
    { if (x[t-d] < r || x[t-d]==r )
      { u=1
        while(u>0)
        { for(j in 1:p1)
          { y[t,j]=rbinom(1,x[t+p1-j],alpha1[j]) }
          if (sum(y[t,])<= x[t+p1])
          { z[t]=x[t+p1]-sum(y[t,])
            u=-1 }}}}
        else
        { u=1
          while(u>0)
          { for(j in 1:p2)
            { y[t,j]=rbinom(1,x[t+p2-j],alpha2[j]) }
            if (sum(y[t,])<= x[t+p2])
            { z[t]=x[t+p2]-sum(y[t,])
              u=-1 }}}}
        for(t in (d+1):n)
        {if (x[t-d] < r || x[t-d]==r )
          {somay1[m]<-somay1[m]+y[t,1]
            somax1[m]<-somax1[m]+x[t+p1-1]
            somaz1[m]<-somaz1[m]+z[t]
            somaum1[m]<-somaum1[m]+1
          }}
        for(t in (d+1):n)
        {if (x[t-d]>r )
          {somay21[m]<-somay21[m]+y[t,1]
            somay22[m]<-somay22[m]+y[t,2]
            somax21[m]<-somax21[m]+x[t+p2-1]
            somax22[m]<-somax22[m]+x[t+p2-2]
            somaz2[m]<-somaz2[m]+z[t]
            somaum2[m]<-somaum2[m]+1
          }}
        m<-m+1 }
      alpha1[1]<- sum(somay1)/sum(somax1)
      alpha2[1]<- sum(somay21)/sum(somax21)
      alpha2[2]<- sum(somay22)/sum(somax22)
      lambda_1<-sum(somaz1)/sum(somaum1)
      lambda_2<-sum(somaz2)/sum(somaum2)
      cont<-cont+1 }
    return(rbind(x,y,prop))}
#####

```

```
##### Simular um conjunto de dados PSETINAR(2;1,1)_T #####
simula_psetinar<-function(T,alpha1,alpha2,lambda,R,n)
{N<-100*T+n;
x<-rep(0,N+T);
x[1]<-round(lambda[1],0);
x[2]<-round(lambda[2],0);
m<-100+trunc(n/T)+1;

aux<-rep(r,100+trunc(n/T)+1);
r_aux<-rep(0,T*m+1);
r_aux[2:(T*m+1)]<-aux;

aux<-rep(lambda,100+trunc(n/T)+1);
lam_aux<-rep(0,T*m+1);
lam_aux[2:(T*m+1)]<-aux;

aux<-rep(alpha1,100+trunc(n/T)+1);
alpha1_aux<-rep(0,T*m+1);
alpha1_aux[2:(T*m+1)]<-aux;

aux<-rep(alpha2,100+trunc(n/T)+1);
alpha2_aux<-rep(0,T*m+1);
alpha2_aux[2:(T*m+1)]<-aux;

for(i in (T+1):(N+1) )
{ if(x[i-T]>r_aux[i-1])
{ x[i]<-rbinom(1,x[i-1],alpha2_aux[i-1])+rpois(1,lam_aux[i-1])}
else{ x[i]<-rbinom(1,x[i-1],alpha1_aux[i-1])+rpois(1,lam_aux[i-1])}}
resp=x[(100*T+1):(N+T)]
resp
}
```

```
#####
```

Apêndice

C

Dados de Manchas Solares

C.1 Aplicação 1

Tabela C.1: Número da manchas solares de Wölfer - anual

Anos										
(1770)	101	82	66	35	31	7	20	92	154	125
(1780)	85	68	38	23	10	24	83	132	131	118
(1790)	90	67	60	47	41	21	16	6	4	7
(1800)	14	34	45	43	48	42	28	10	8	2
(1810)	0	1	5	12	14	35	46	41	30	24
(1820)	16	7	4	2	8	17	36	50	62	67
(1830)	71	48	28	8	13	57	122	138	103	86
(1840)	63	37	24	11	15	40	62	98	124	96
(1850)	66	64	54	39	21	7	4	23	55	94
(1860)	96	77	59	44	47	30	16	7	37	74

3.2 Aplicação 2

Tabela 3.2: Série temporal do número total de grupos de manchas solares no período de 1 de janeiro de 1990 a 31 de dezembro de 1991 - Centro nacional de dados geofísicos de Boulder (Colorado, EUA)

Semanas														
(1)	12	15	21	25	18	12	12	20	14	7	13	17	10	15
(15)	11	12	15	9	9	15	13	13	15	12	11	19	15	10
(29)	9	20	19	19	23	24	15	17	18	12	11	17	16	16
(43)	14	13	19	9	8	13	25	14	4	8	15	11	10	15
(57)	18	16	20	15	14	11	13	15	12	16	18	10	19	11
(71)	12	14	16	15	17	21	19	20	19	15	9	17	18	12
(85)	23	30	23	18	13	15	13	10	21	15	12	11	14	8
(99)	12	16	14	17	14	8								

Tabela 3.3: Série temporal do número total de grupos de manchas solares no período de 1 de janeiro de 1990 a 31 de dezembro de 1991 - Palehua Solar observatory (Hawaii, EUA)

Semanas														
(1)	12	17	13	0	18	14	19	17	0	11	20	23	17	17
(15)	16	21	13	16	16	17	15	16	17	14	15	22	17	16
(29)	20	21	20	26	25	24	17	20	20	22	16	22	20	21
(43)	17	21	18	6	10	23	25	14	5	15	19	17	15	16
(57)	17	17	26	23	23	17	12	5	12	19	26	16	18	17
(71)	14	18	19	19	21	23	19	25	22	22	18	19	26	17
(85)	26	28	22	19	15	20	18	17	22	16	15	16	22	10
(99)	18	22	21	18	16	9								

Referências Bibliográficas

- Ahn, S., Gyemin, L. & Jongwoo, J. (2000). Analysis of the M/D/1-type queue based on the integer-valued autoregressive process. *Oper. Res. Lett.* **27**, 235-241.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proc. 2nd Inter. Symposium on Information Theory*, pp. 267–281, Budapest.
- Al-Osh, M.A. & Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **8**, 261-275.
- Al-Osh, M.A. & Alzaid, A.A. (1988). Integer-valued moving average (INMA) process. *Stat. Papers* **28**, 281-300.
- Al-Osh, M.A. & Alzaid, A.A. (1991). Binomial autoregressive moving average models. *Stoch. Models* **7**, 261-282.
- Al-Osh, M.A. & Aly, E.E.A.A. (1992). First order autoregressive time series with negative binomial and geometric marginals. *Commun. Statist. Theory Math.* **21**, 2483-2492.
- Alzaid, A.A. & Al-Osh, M.A. (1990). An integer-valued p th order autoregressive structure (INAR(p)) process. *J. Appl. Prob.* **27**, 314-324.

- Alzaid, A.A. & Al-Osh, M.A. (1993). Some autoregressive moving average processes with generalized Poisson marginal distributions. *Ann. Inst. Statist. Math.* **45**, 223-232.
- Anderson, J. & Karlis, D. (2010). Treating missing values in INAR(1) models: An application to syndromic surveillance data. *J. Time Ser. Anal.* **31**, 12-30.
- Bennett, W.R. (1958). Statistics of regenerative digital transmission. *Bell Sys. Tech. J.* **37**, 1501-1542.
- Berglund, E. & Brännäs, K. (1996). Entry and exit of plants: A study based on Swedish panel count data. In *1995 Yearbook of the Finnish Statistical Society*, Helsinki, 95–111.
- Besag, J. & Green, P.J. (1993). Spatial statistics and Bayesian computation. *J. Royal Statist. Soc. B* **55**, 25-37.
- Best, N.G., Cowles, M.K. & Vines, K. (1995). Coda: Convergence diagnosis and output analysis software for gibbs sampling output. Version 0.3. *Technical report*, Biostatistics Unit MRC, Cambridge, Inglaterra.
- Billingsley, P. (1961). *Statistical Inference for Markov Processes*. University of Chicago Press, Chicago.
- Box, G.E.P. & Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.
- Box, G.E.P., Jenkins, G. & Reinsel, C. (2008). *Time Series Analysis: Forecasting and Control*. Wiley, 4th edition.
- Brännäs, K. (1994). Estimation and testing in integer-valued AR(1) models. *Umeå Economic Studies*, 335.
- Brännäs, K. (1995). Explanatory variables in the AR(1) Poisson model. *Umeå Economic Studies*, 381.

- Brännäs, K. & Hellström, J. (2001). Generalized integer-valued autoregression. *Econometric Rev.* **20**, 425–443.
- Brännäs, K., Hellström, J. & Nordström, J. (2002). A new approach to modelling and forecasting monthly guest nights in hotel. *Int. J. Forecast.* **18**, 9-30.
- Brännäs, K. & Nordström, J. (2006). Tourist accommodation effects of festivals. *Tourism Econ.* **12**, 291-302.
- Brännäs, K. & Quoreshi, A.M.M.S. (2010). Integer-valued moving average modelling of the number of transactions in stocks. *Appl. Finan. Econ.* **20**, 1429–1440.
- Brooks, S.P., Giudici, P. & Roberts, G.O. (2003). Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions. *J. Royal Statist. Soc. B* **65**, 3-55.
- Bu, R., McCabe, B. & Hadri, K. (2008). Maximum likelihood estimation of higher-order integer-valued autoregressive processes. *J. Time Ser. Anal.* **6**, 973–994.
- Campbell, E.P. (2004). Bayesian selection of threshold autoregressive models. *J. Time Ser. Anal.* **25**, 467-482.
- Casella, G. & Berger, R.L. (2001). *Statistical Inference*. Brooks/Cole, 2nd edition.
- Chan, K.S. (1993). Asymptotic behaviour of the Gibbs sampler. *J. Amer. Statist. Assoc.* **88**, 320-326.
- Chen, M.-H. & Ibrahim, J.G. (2000). Bayesian predictive inference for time series count data. *Biometrics* **56**, 678–685.
- Dempster, A.P., Laird, N.M. & Rubin, D. B. (1985). Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statist. Soc. B* **39**, 1-38.
- Dion, J-P., Gauthier, G. & Latour, A. (1995). Branching processes with immigration and integer-valued time series. *Serdica Math. J.* **21**, 123–136.

- Du, J.-Guan & Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* **12**, 129–142.
- Durbin, J. & Koopman, S.J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives. *J. Royal Statist. Soc. B* **62**, 3–56.
- Dyk, D.A. van & Meng, X. (2001). The art of data augmentation. *J. Comp. Graph. Stat.* **10-1**, 1–50.
- Enciso-Mora, V., Neal, P.J. & Subba Rao, T. (2008). Efficient order selection algorithms for integer-valued ARMA processes. *J. Time Ser. Anal.* **30**, 1–18.
- Enciso-Mora, V., Neal, P.J. & Subba Rao, T. (2009). Integer valued AR processes with explanatory variables. *Sankhya, The Indian Journal of Statistics* **71-B**, 248–263.
- Freeland, R.K. (1998). Statistical analysis of discrete time series with application to the analysis of workers compensation claims data. *Unpublished doctoral dissertation*. The University British Columbia, Canada.
- Franke, J. & Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modeling epileptic seizure counts. *Time Series Analysis* (T. Subba Rao, ed.). London, Chapman and Hall, pp. 310–330.
- Franke, J. & Subba Rao, T. (1995). Multivariate first order integer valued autoregressions. *Technical Report*, Universitat Kaiserslautern.
- Gamerman, D. (1996). Simulação Estocástica via Cadeias de Markov. *Instituto de Matemática UFRJ*.
- Geman, S. & Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 721–741.

- Gelfand, A.E. & Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85**, 398-409.
- Gelman, A. & Rubin, D.B. (1992). A single series from the Gibbs sampler provides a false sense of security. *Bayesian Statistics* (J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, eds.). Oxford University Press **4**, pp. 625-631.
- Gilks, W.R. & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *J. Royal Statist. Soc. C* **41**, 337-348.
- Gilks, W.R., Richardson, S. & Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall.
- Gladyshev, E.G. (1961). Periodically correlated random sequences. *Soviet Math* **2**, 385-388.
- Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711-732.
- Gomes, D. & Canto e Castro, L. (2009). Generalized integer-valued random coefficient for a first structure autoregressive (RCINAR). *J. Statis. Plann. Inference* **139**, 4088-4097.
- Hastings, W.K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* **57**, 97-100.
- Hoyt, D.V. & Schatten, K. (1998). Group sunspot numbers: A new solar activity reconstruction. *Solar Phys.* **179**, 189-219.
- Izenman, A.J., Wolf, J. R. & Wolfer, H.A. (1983). An historical note on the Zurich sunspot relative numbers. *J. Royal Statist. Soc. A* **146**, 311-318.
- Jefreys, H. (1998). *Theory of Probability*. Oxford University Press, New York.

- Jones, R.H. & Brelsford, W.M. (1967). Time series with periodic structure. *Biometrika* **54**, 403-408.
- Keith, J.M., Kroese, D.P. & Sofronov, G. (2008). Adaptive independence samplers. *Stat. Comput.* **18**, 409-420.
- Klimko, L.A. & Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* **6**, 629-642.
- Latour, A. (1997). The multivariate GINAR(p) process. *Adv. Appl. Probab.* **29**, 228-248.
- Latour, A. (1998). Existence and stochastic structure of a non negative integer-valued autoregressive process. *J. Time Ser. Anal.* **4**, 439-455.
- Man-Suk Oh & Bin Lim, Y. (2001). Bayesian analysis of time series Poisson data. *J. Appl. Statist.* **28**, 259-271.
- McCabe, B.P.M. & Martin, G.M. (2005). Bayesian predictions of low count time series. *Int. J. Forecast.* **21**, 315-330.
- McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Res. Bull.* **21**, 645-650.
- McKenzie, E. (1986). Autoregressive-moving average process with negative-binomial and geometric marginal distributions. *Adv. Appl. Probab.* **18**, 679-705.
- McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Probab.* **20**, 822-835.
- McKenzie, E. (2003). Discrete variate time series. *Handbook of Statistics*. (C.R. Rao and D.N. Shanbhag, eds.). Amsterdam, Elsevier, pp. 573-606.
- McLachlan, G. & Krishnan, T. (1997). *The EM Algorithm and Extensions*. Wiley Series in Probability and Statistics. John Wiley & Sons.

- Metropolis, N., Rosenblut, A.W., Rosenblut, M.N. & Teller, E. (1953). Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087-1092.
- Monteiro, M. (2010). Contributos para a Análise Estatística de Séries de Contagem. *Unpublished doctoral dissertation*, Universidade de Aveiro, Portugal.
- Monteiro, M., Pereira, I. & Scotto, M.G. (2008). Optimal alarm system for count process. *Comm. Statist. Theory Methods* **37**, 3054-3076.
- Monteiro, M., Scotto, M.G. & Pereira, I. (2010). Integer-valued autoregressive processes with periodic structure. *J. Statist. Plann. Inference* **140**, 1529–1541.
- Monteiro, M., Scotto, M.G. & Pereira, I. (2012). Integer-valued self-exciting threshold autoregressive processes. *Comm. Statist. Theory Methods* **41**, 2717–2737.
- Neal, R.M. (2003). Slice sampling. *Ann. Statist.* **31**, 705-767.
- Neal, P.J. & Subba Rao, T. (2007). MCMC for integer valued ARMA processes. *J. Time Ser. Anal.* **28**, 92-110.
- Nordström, J. (1996). Tourism satellite account for Sweden 1992-93. *Tourism Econ.* **2**, 13-42.
- Oliveira, V.L. (2000). Modelos autoregressivos para sucessões cronológicas de contagem: caracterização e modelação. *Unpublished doctoral dissertation*, Universidade do Porto, Portugal.
- Paulino, C.D., Turkman, A.A. & Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian. Lisboa.
- Pavlopoulos, H. & Karlis, D. (2008). INAR(1) modeling of overdispersed count series with an environmental application. *Environmetrics* **19**, 369–393.
- Quoreshi, A.M.M.S. (2006). A vector integer-valued moving average model for high frequency financial count data. *Econ. Letters* **101**, 258-261.

- Quoreshi, A.M.M.S. (2012). A long-memory integer-valued time series model INAR-FIMA, for financial application. *Quant. Finance* (to appear).
- Ritter, C. & Tanner, M. (1992). Facilitating the Gibbs sampler: The Gibbs stopper and the Griddy-Gibbs sampler. *J. Amer. Statist. Assoc.* **87**, 861-868.
- Roberts, G.O. (1996). Markov chain concepts related to sampling algorithms. *Markov Chain Monte Carlo in Practice*. (W.R. Gilks, S. Richardson and D.J. Spiegelhalter, eds.). Chapman and Hall, pp. 45-57.
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.* **6**, 461-464.
- Scotto, M.G., Weiß, C.H., Silva, M.E. & Pereira, I. (2014). Bivariate binomial autoregressive models. *J. Multivariate Anal.* **125**, 233-251.
- Silva, I. (2005). Contributions to the analysis of discrete-valued time series. *Unpublished doctoral dissertation*, Universidade do Porto, Portugal.
- Silva, I. & Silva, M.E. (2003). Seleção de ordem em modelos INAR. *Literacia e Estatística (Actas do X Congresso Anual da SPE)*. (P. Brito, A. Figueiredo, F. Sousa, P. Teles e F. Rosado, eds.). Edições SPE, pp. 627-638.
- Silva, I. & Silva, M.E. (2004). Estimação de Whittle para réplicas de processos INAR. *Estatística com Acaso e Necessidade (Actas do XI Congresso Anual da SPE)*. (P. Rodrigues, E. Rebelo e F. Rosado, eds.). Edições SPE, pp. 767-778.
- Silva, I., Silva, M.E., Pereira, I. & Silva, N. (2005). Replicated INAR(1) processes. *Methodol. Comput. Appl. Probab.* **7**, 517-542.
- Silva, M.E. & Oliveira, V. L. (2004). Difference equations for the higher-order moments and cumulants of the INAR(1) model. *J. Time Ser. Anal.* **25**, 317-333.
- Silva, M.E. & Oliveira, V. L. (2005). Difference equations for the higher-order moments and cumulants of the INAR(p) model. *J. Time Ser. Anal.* **26**, 17-36.

- Silva, N. (2005b). Análise bayesiana de séries temporais de valores inteiros. *Unpublished doctoral dissertation*, Universidade de Aveiro, Portugal.
- Smith, B.J. (2004). Bayesian Output Analysis Program (BOA) for MCMC. R package version 1.1.2-1, URL <http://www.public-health.uiowa.edu/boa>.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *J. Royal Statist. Soc. B* **64**, 583–639.
- Steutel, F. W. & van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Probab.* **5**, 893–899.
- Stramer, O. & Lin, Y. (2002). On inference for threshold autoregressive models. *Test* **11**, 51–71.
- Tanner, M.A. & Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82**, 528–550.
- Tanner, M.A. (1996). *Tools for Statistical Inference*. Springer-Verlag, 3rd edition, 208.
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22**, 1701–1762.
- Thyregod, P., Carstensen, J., Madsen, E. & Arnbjerg-Nielsen, K. (1999). Integer valued autoregressive models for tipping bucket rainfall measurements. *Environmetrics* **10**, 395–411.
- Tong, H. (1983). *Threshold Models in Non-linear Time Series Analysis*. Springer-Verlag Inc, New York.
- Vermaak, J., Andrieu, C., Doucet, A. & Godsill, S.J. (2004). Reversible jump Markov chain Monte Carlo strategies for Bayesian model selection in autoregressive processes. *J. Time Ser. Anal.* **25**, 785–809.

- Wei, W.S. (1989). *Time Series Analysis Univariate and Multivariate Methods*. Redwood City, Addison Wesley.
- Wei, G.C.G. & Tanner, M.A. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *J. Amer. Statist. Assoc.* **85**, 699-704.
- Wei, C.H. (2008a). Thinning operations for modeling time series of counts a survey. *Adv. Statist. Anal.* **92**, 319-341.
- Wei, C.H. (2008b). The combined INAR(p) models for time series of counts. *Statist. Probab. Lett.* **78**, 1817-1822.
- West, H., Harrison, P.J. & Migon, H.S. (1985). Dynamic generalized linear models and bayesian forecasting. *J. Amer. Statist. Assoc.* **80**, 73-83.
- Wu, C.F.J. (1983). On the convergence properties of the EM algorithm. *Ann. Statist.* **11**, 95-103.
- Zhou, R. & Basawa, I.V. (2005). Least-squares estimation for bifurcation autoregressive process. *Statis. Probab. Lett.* **74**, 77-88.